UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

DANIEL ALVES DA SILVA LOPES DINIZ

PROBABILIDADE DO CONDICIONAL E PROBABILIDADE CONDICIONAL:

LEIS DO EXCESSO E RESULTADOS DE TRIVIALIZAÇÃO NA LÓGICA LFI1

CAMPINAS

2020

DANIEL ALVES DA SILVA LOPES DINIZ

PROBABILIDADE DO CONDICIONAL E PROBABILIDADE CONDICIONAL: LEIS DO EXCESSO E RESULTADOS DE TRIVIALIZAÇÃO NA LÓGICA LFI1

Dissertação apresentada ao Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Filosofia.

Orientadora: Profª. Drª. Juliana Bueno
Coorientador: Prof. Dr. Walter Alexandre Carnielli

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO DANIEL ALVES DA SILVA LOPES DINIZ, ORIENTADA POR PROFª. DRª. JULIANA BUENO.

CAMPINAS
2020

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Filosofia e Ciências Humanas
Paulo Roberto de Oliveira - CRB 8/6272

D615p

Diniz, Daniel Alves da Silva Lopes, 1995-
Probabilidade do condicional e probabilidade condicional : leis do excesso e resultados de trivialização na lógica LFI1 / Daniel Alves da Silva Lopes Diniz. – Campinas, SP : [s.n.], 2020.

Orientador: Juliana Bueno.
Coorientador: Walter Alexandre Carnielli.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas.

1. Lógica paraconsistente. 2. Probabilidades - Filosofia. 3. inconsistência (Lógica). I. Bueno, Juliana, 1976-. II. Carnielli, Walter Alexandre, 1952-. III. Universidade Estadual de Campinas. Instituto de Filosofia e Ciências Humanas. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Probability of the conditional and conditional probability : excess laws and triviality results in the LFI1 logic
Palavras-chave em inglês:
Paraconsistent logic
Probabilities - Philosophy
Inconsistency (Logic)
Área de concentração: Filosofia
Titulação: Mestre em Filosofia
Banca examinadora:
Juliana Bueno [Orientador]
Julio Michael Stern
Pedro Merlussi
Data de defesa: 27-01-2020
Programa de Pós-Graduação: Filosofia

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0001-8874-3084
- Currículo Lattes do autor: http://lattes.cnpq.br/7560668444401188

UNICAMP

Universidade Estadual de Campinas
Instituto de Filosofia e Ciências Humanas

A Comissão julgadora dos trabalhos de Defesa de dissertação de Mestrado, composta pelos Doutores a seguir descritos em sessão pública realizada em 27 de janeiro de 2020, considerou o candidato Daniel Alves da Silva Lopes Diniz aprovado.

Profª. Drª. Juliana Bueno

Prof. Dr. Julio Michael Stern

Dr. Pedro Merlussi

A Ata de Defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertações/Teses e na Secretaria do Programa de Pós-Graduação em Filosofia do Instituto de Filosofia e Ciências Humanas.

Para meu pai, por ser a minha família.

Agradecimentos

À professora Juliana e ao professor Carnielli, pela orientação. À secretária do PPG de Filosofia, Daniela, e aos funcionários do Centro de Lógica, Epistemologia e História da Ciência da Unicamp, pela paciência. À Fernanda Caldas, por ter revisado o Abstract.

A versão desta dissertação submetida ao exame de qualificação estava bastante confusa, e algumas partes provavelmente quase ilegíveis para qualquer um que não eu mesmo. Nesse sentido, agradeço a paciência dos professores Abílio e Evandro em lerem aquele texto incipiente e apontarem várias possibilidades de melhorias. Espero ter conseguido implementar pelo menos as mais cruciais. Agradeço também a disposição em aceitar meu convite tão apressado e em cima da hora para participarem da banca.

De modo semelhante, o texto que submeti à defesa continha ainda alguns erros importantes e o fluxo de leitura entre suas seções era prejudicado por interrupções e cortes bruscos de raciocínio. Agradeço ao doutor Merlussi e ao professor Stern pela paciência em, ainda assim, ler tudo e recomendar correções. Agradeço ainda pela participação da professora Matulovic na banca, mesmo sua presença não sendo obrigatória. Qualquer erro ou omissão que tenha eventualmente restado na versão final do texto é responsabilidade exclusivamente minha.

Ao meu pai, por me apoiar, acalmar e reconfortar ao longo de todo o processo de escrita. Em tudo que faço há uma tentativa de retribuir seu amor. Aos meus amigos, que tornaram muito mais tranquila e prazerosa a redação desta dissertação. Ao Felipe e à Mari, por todos os momentos! Ao Grupo de Lógica Matemática do IMECC, por tantas jantas interessantes falando sobre Lógica. Ao Grupo de Escalada Esportiva da Unicamp, pelo acolhimento e incentivo. Daqui a pouco volto a escalar. . . ! À Fátima e ao José Luiz, por acompanharem de perto minha vida e, em particular, a sofrida escrita desta dissertação.

A todas as pessoas empenhadas em democratizar o acesso ao conhecimento científico e atenuar ou remover empecilhos injustos à prática acadêmica.

Às inúmeras pessoas que desenvolvem *softwares* livres: a redação desta dissertação foi imensamente facilitada por essas iniciativas. Este texto foi formatado em LaTeX, o sistema tipográfico TeX de Donald Knuth enriquecido pelas macros de Leslie Lamport. Os arquivos fonte foram escritos em vários editores de texto livres na distribuição Ubuntu de GNU/Linux, e compilados por pdfTeX, desenvolvido por Hàn Thế Thành. As várias etapas do processo de escrita foram arquivadas com o sistema de controle de versões Apache Subversion.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Código de Financiamento 001. São meus votos que as agências de fomento à pesquisa e as universidades brasileiras recuperem-se o mais cedo possível dos ataques infligidos pelo governo federal atualmente no poder.

Averiguar o distante

sem sair de casa	conhecer o mundo
sem espiar pela janela	ver o curso do céu
quanto mais longe se vai	tanto menos se conhece

por isso o homem santo. . .

não anda. . .	e conhece
não vê. . .	e nomeia
não atua. . .	e realiza

Lao Zi, *Dao De Jing*.
Organização e tradução de Mario Sproviero.
São Paulo: Hedra, 2014.

Resumo

Desde cedo nas investigações sobre a viabilidade da incorporação da teoria da probabilidade à Lógica, percebeu-se que expressões condicionais da linguagem natural têm duas formalizações intuitivas. Sejam α, β proposições e $0 \leq x \leq 1$ um valor de probabilidade. Considere-se, então, a sentença "a probabilidade de α dada β é x". Por um lado, ela parece denotar uma proposição condicional; por outro, parece formalizável como uma proposição expressando uma probabilidade condicional. Uma solução simples para essa ambiguidade interpretativa é estipular que $p(\alpha \to \beta)$ é igual não apenas a $p({\sim}\alpha \vee \beta)$, mas também a $p(\beta \mid \alpha)$. Karl R. Popper parece ter sido o primeiro a perceber que, em geral, $p(\alpha \to \beta) > p(\beta \mid \alpha)$, e que $p(\alpha \to \beta) = p(\beta \mid \alpha)$ se, e somente se, $p({\sim}\alpha) = 0$ ou $p(\beta \mid \alpha) = 1$. Esses resultados ficaram conhecidos como "leis do excesso". Pensou-se, então, que substituir a implicação material por outro condicional restauraria a igualdade intuitiva entre os valores. Lewis (1976) provou, no entanto, que isso não é o caso: se $\rightsquigarrow$ é um condicional arbitrário definido tal que $p(\alpha \rightsquigarrow \beta) := p(\beta \mid \alpha)$, então $p(\alpha \rightsquigarrow \beta) = p(\beta)$, isto é, α e β são proposições independentes, o que é absurdo, dado que podem ser quaisquer duas proposições. Outros teoremas equivalentes foram demonstrados desde então; os mais fortes dessa família encontram-se em Fitelson (2015). Posto isso, essa pesquisa objetiva responder a seguinte questão: as leis do excesso e os resultados de trivialização de Lewis (1976) *et al.* são válidos caso a lógica à qual se incorpora a teoria da probabilidade seja paraconsistente? Se são, em que medida a demonstração dessa validade é análoga às demonstrações originais? Se não são válidos, como se prova que não o são? Dentre as lógicas paraconsistentes, aquelas em que não vigora o princípio segundo o qual $\{\alpha, {\sim}\alpha\} \vdash \beta$, para quaisquer proposições α e β, destacam-se as lógicas da inconsistência formal (LFIs). Elas internalizam o conceito metalógico de consistência em suas linguagens-objeto, definindo um conectivo primitivo unário $\circ$ tal que $\circ\alpha$ denota que α é uma proposição consistente. LFI1 (também conhecida como J3) é a LFI mais próxima da lógica proposicional clássica. Assim, pretende-se aqui investigar, mais especificamente, o *status* das leis do excesso e dos resultados de trivialização no fragmento estritamente inconsistente de LFI1. Embora não se tenha conseguido provar que esses resultados são indemonstráveis em LFI1, constatou-se que as demonstrações originais não podem ser reproduzidas devido às propriedades formais da negação e da implicação paraconsistentes. A análise desse fato é bastante valiosa em si mesma, pois evidencia várias características metalógicas importantes de LFI1 (e das LFIs e das lógicas paraconsistentes em geral), e avança o debate em torno da articulação entre Lógica e teoria da probabilidade, que afigura-se muito promissora quanto a seus ganhos em poder expressivo.

Palavras-chave: lógica paraconsistente, Probabilidades—Filosofia, inconsistência (Lógica).

Abstract

Since early in the investigations on the viability of incorporating probability theory into Logic, it has been noticed that conditional natural language expressions have two intuitive formalizations. Let α, β be propositions and $0 \leq x \leq 1$ a probability value. Consider, then, the sentence "the probability of α given β is x". On the one hand, it seems to denote a conditional proposition; on the other hand, it seems formalizable as a proposition expressing a conditional probability. A simple solution to this interpretative ambiguity is to estipulate that $p(\alpha \to \beta)$ equals not only $p({\sim}\alpha \vee \beta)$, but also $p(\beta \mid \alpha)$. Karl R. Popper seems to be the first to have noticed that, in general, $p(\alpha \to \beta) > p(\beta \mid \alpha)$, and that $p(\alpha \to \beta) = p(\beta \mid \alpha)$ if, and only if, $p({\sim}\alpha) = 0$ or $p(\beta \mid \alpha) = 1$. These results have become known as "excess laws". It was thought, then, that replacing the material implication with another conditional would restore the intuitive equality between the values. Lewis (1976) proved, however, that this is not the case: if $\rightsquigarrow$ is an arbitrary conditional defined such that $p(\alpha \rightsquigarrow \beta) := p(\beta \mid \alpha)$, then $p(\alpha \rightsquigarrow \beta) = p(\beta)$, that is, α and β are independent propositions, which is absurd, since they may be any two propositions. Other equivalent theorems have been proven since then; the strongest ones of this familiy are to be found in Fitelson (2015). That said, this research aims to answer the following question: are the excess laws and the trivialization results of Lewis (1976) *et al.* valid if the logic to which probability theory is incorporated into is paraconsistent? If so, to which degree the proof of this validity is analogous to the original proofs? Otherwise, how could it be proven that they are not? Among paraconsistent logics, those in which the principle that $\{\alpha, {\sim}\alpha\} \vdash \beta$, for any propositions α and β, is not valid, the logics of formal inconsistency (LFIs) stand out. They internalize the metalogical concept of consistency in their object-languages, defining an unary primitive conective $\circ$ such that $\circ\alpha$ denotes that α is a consistent proposition. LFI1 (also known as J3) is the closest LFI to propositional classical logic. Therefore, it is intended here, more specifically, to investigate the status of the excess laws and the trivialization results in the strictly inconsistent fragment of LFI1. Although it was not possible to prove that these results cannot be proven in LFI1, it has been noticed that the original proofs cannot be replicated due to the formal properties of paraconsistent negation and implication. The analysis of this impossibility is very valuable itself, because it puts into evidence many important metalogical characteristics of LFI1 (and of LFIs and paraconsistent logics in general), and advances the debate on the articulation between Logic and probability theory, which stands as very promising as to its gains in expressive power.

Keywords: paraconsistent logic, probabilities—Philosophy, inconsistency (Logic).

Lista de abreviaturas e siglas

fbf	fórmula bem formada
FND	forma normal disjuntiva
LPC	lógica proposicional clássica
SQL	linguagem de consulta estruturada (*Structured Query Language*)

Lista de símbolos

Metalinguagem

$\Longleftrightarrow$	se, e somente se (metalinguagem)

Teoria de conjuntos

$\{\ldots\}$	conjunto
$\{x\colon\ \ldots(x)\}$	conjunto dos objetos que satisfazem um predicado
$\mathrm{A}, \mathrm{B}, \Gamma, \ldots$	conjuntos
$\{\alpha_\beta\}_{\beta\in\mathrm{B}}$	conjunto indexado por um conjunto B
$\lvert\ldots\rvert$	cardinalidade
$\in$	pertence
$\notin$	não pertence
$\subseteq$	está contido
$\nsubseteq$	não está contido
$\cup$	união
$\bigcup \mathrm{A}$	união de todos os subconjuntos de A
$\bigcup_{i=0}^{k} \mathrm{A}_i$	$\mathrm{A}_0 \cup \ldots \cup \mathrm{A}_k$
$\cap$	intersecção
$\times$	produto cartesiano
A^n	$\mathrm{A} \times \mathrm{A} \times \ldots \times \mathrm{A}$ (n vezes)
$\setminus$	diferença (de conjuntos)
$\varnothing$	conjunto vazio
$\wp(\ldots)$	conjunto das partes
Ω	espaço amostral (de probabilidade)
$\mathfrak{F}$	álgebra de conjuntos
ω	primeiro ordinal infinito
$\mathbb{N}$	conjunto dos números naturais
$\mathbb{N}^+$	conjunto dos números naturais positivos
$\mathbb{R}$	conjunto dos números reais
$\langle\ldots\rangle$	ênupla
$\langle x_1, \ldots, x_n\rangle[i]$	x_i, se $1 \leq i \leq n$

Lógica

$f, g, h, \ldots$	funções
$f\colon \mathrm{A} \mapsto \mathrm{B}$	função de domínio A e contradomínio B
f^{-1}	função inversa de f

$\mathcal{L}_{\mathrm{L}}$	linguagem de uma lógica L
$\mathrm{P}, \mathrm{Q}, \mathrm{R}, \ldots$	predicados de lógica de primeira ordem
$\exists$	quantificador existencial (existe)
$\forall$	quantificador universal (para todo)
$\nexists$	não existe
$a, b, c, \ldots$	variáveis proposicionais
$\alpha, \beta, \gamma, \ldots$	metavariáveis proposicionais
$\mathcal{P}$	conjunto enumerável de variáveis proposicionais
$\top$	partícula *top*
$\bot$	partícula *bottom*
$\mathtt{t}$	verdade
$\mathtt{f}$	falsidade
$\mathbf{2}$	o conjunto $\{\mathtt{t}, \mathtt{f}\}$
$\neg$	negação (fraca)
$\sim$	negação forte
$\vee$	disjunção
$\bigvee_{i=0}^{k} \alpha_i$	$\alpha_0 \vee \ldots \vee \alpha_k$
$\wedge$	conjunção
$\rightarrow$	implicação material
$\leftrightarrow$	equivalência material
$\rightsquigarrow$	condicional genérico
∇	possibilidade
$\bullet$	inconsistência
$\circ$	consistência
$=$	igualdade
$:=$	definição
Σ	assinatura
$\vdash$	dedução sintática
$\dashv\vdash$	dedução sintática recíproca
$\models$	dedução semântica
$\not\models$	dedução semântica negada
$=\mid\models$	dedução semântica recíproca
	Matemática
$\approx$	igualdade aproximada
$\cdot$	produto
$[i, j]$	intervalo fechado entre i e j
(i, j)	intervalo aberto entre i e j
$\mathrm{A}, \mathrm{B}, \mathrm{C}, \ldots$	pontos

$\widehat{\mathrm{AB}}$	arco entre os pontos A e B
$\overline{\mathrm{AB}}$	segmento de reta cujas extremidades são A e B
$\triangle\mathrm{ABC}$	triângulo de vértices $\mathrm{A}, \mathrm{B}, \mathrm{C}$

SUMÁRIO

Introdução

> Do not conclude from your apprenticeship that nothing is left for you to know, but that there is an infinity left for you to know.
>
> Pascal, *Pensées*. Tradução de Gertrude Burfurd Rawlings.

O objetivo original desta dissertação é avaliar se certos resultados particularmente interessantes relativos à interface entre lógica[1] clássica e teoria axiomática de probabilidade vigoram também na lógica paraconsistente LFI1. Se vigoram, de que maneira suas demonstrações diferem das demonstrações existentes na literatura para o caso clássico? Se não vigoram, como é possível comprovar esse fato? O caso negativo é mais espinhoso, porque é sempre mais complexo demonstrar uma impossibilidade que uma possibilidade.

Os resultados mencionados concernem a expectativa bastante intuitiva de que o valor de probabilidade absoluta de uma implicação material é igual ao respectivo valor de probabilidade condicional (ou relativa). Isto é, tende-se a supor que $p(\alpha \rightarrow \beta) = p(\alpha \mid \beta)$[2], e, com efeito, essa posição foi explícita ou implicitamente adotada com frequência no início dos estudos sobre a interface entre lógica (clássica) e teoria da probabilidade. Os resultados analisados aqui demonstram, justamente, que essa igualdade não se verifica. Mais ainda, não se verifica mesmo que se substitua a implicação material por um conectivo condicional genérico, definido especialmente para que seu valor de probabilidade absoluta iguale-se ao valor de probabilidade condicional correspondente. Isto é, mesmo a igualdade $p(\alpha \rightsquigarrow \beta) = p(\alpha \mid \beta)$ não é o caso. Assim, a meta inicial desta dissertação é verificar se essa desigualdade entre $p(\alpha \rightsquigarrow \beta)$ e $p(\alpha \mid \beta)$ vale também na lógica paraconsistente LFI1.

Além desse objetivo principal específico e "técnico", considerou-se necessário, antes, situar minimamente o leitor em teoria da probabilidade, e em Lógica, especialmente em lógicas paraconsistentes. Esses temas, embora relacionados apenas indiretamente à questão central desta dissertação, são pertinentes aqui porque uma visão panorâmica adequada deles proporciona uma compreensão mais fácil e completa das discussões específicas deste trabalho. Além disso, a teoria da probabilidade não é um assunto comumente abordado nos cursos de Filosofia brasileiros, tanto de graduação quanto de pós-graduação, de modo que as pessoas versadas em Filosofia e Lógica, que são, afinal, o público-alvo deste trabalho, beneficiar-se-ão de uma introdução a esse tópico. As lógicas paraconsistentes também não são familiares à maioria dos graduandos ou graduados

1 Adotar-se-á a seguinte convenção para a escrita da palavra "lógica": quando grafada assim, ela denotará um certo sistema lógico, isto é, um certo conjunto de (esquemas de) axiomas munido de uma ou mais regras de inferência. A grafia "Lógica" será reservada à disciplina que estuda esses sistemas.

2 A notação lógica empregada aqui será propriamente explicada adiante. Por ora, observe-se apenas que a implicação material clássica será denotada por $\rightarrow$, e $\rightsquigarrow$ representa um conectivo condicional genérico.

em Filosofia, motivo pelo qual também são brevemente apresentadas aqui. Os interessados em "ir direto ao ponto" podem pular os capítulos 1, 2 e 4, que são aqueles dedicados a essas questões mais periféricas.

Algumas noções de Teoria de conjuntos serão necessárias imediatamente, de modo que serão apresentadas ainda nesta Introdução. Como o sabe qualquer pessoa que já tenha começado seus estudos em Lógica ou em Teoria de conjuntos, algum conhecimento de uma é sempre necessário para o estudo da outra. Um modo razoável e comum de evitar que essa interdependência torne-se um círculo vicioso é introduzir a área de interesse assumindo apenas noções intuitivas da outra. Adota-se, aqui, essa estratégia: sendo esta uma dissertação sobre Lógica e não sobre Teoria de conjuntos, a primeira será apresentada de modo que serão necessários apenas intuições básicas sobre a última. Em particular, não será preciso conhecer teoria axiomática de conjuntos[3].

Optou-se por uma notação bastante padrão. Um conjunto pode ser representado por:

i) uma letra grega maiúscula $(\mathrm{A}, \mathrm{B}, \Gamma, \ldots)$; ou
ii) uma expressão da forma $\{x\colon \alpha(x)\}$, em que α é uma fórmula lógica na qual x ocorre livremente e cuja satisfação determina o pertencimento ao conjunto; ou
iii) uma listagem de seus elementos separados por vírgulas, delimitada por chaves (por exemplo: $\{a, b, c\}$)[4]; ou
iv) expressões da forma $\{\alpha_\beta\}_{\beta \in \mathrm{B}}$, cujo significado será explicado adiante.

Os símbolos $\varnothing, \mathbb{N}, \mathbb{N}^+, \mathbb{R}$ representam os conjuntos vazio, dos números naturais, dos números naturais positivos e dos números reais, respectivamente. Os símbolos $\in, \subseteq, \subset$ denotam as relações de pertencimento, continência e continência estrita, respectivamente. A está contido em B ($\mathrm{A} \subseteq \mathrm{B}$) se todo elemento que pertence a A pertence a B. A está contido estritamente em B ($\mathrm{A} \subset \mathrm{B}$) se todo elemento que pertence a A pertence a B, mas a recíproca é falsa. $a \notin \mathrm{A}$ abrevia $\sim(a \in \mathrm{A})$, e $\mathrm{A} \nsubseteq \mathrm{B}$ abrevia $\sim(\mathrm{A} \subseteq \mathrm{B})$. $\mathrm{A} \setminus \mathrm{B}$ simboliza a diferença entre A e B, o conjunto $\{a\colon a \in \mathrm{A} \text{ e } a \notin \mathrm{B}\}$. Os símbolos $\cup, \cap$ representam as operações de união e intersecção, respectivamente. $\bigcup \mathrm{A}$ denota o conjunto que contém exatamente todos os elementos de cada elemento de A. A cardinalidade de um conjunto A é representada por $|\mathrm{A}|$. $\wp(\mathrm{A})$ denota o conjunto das partes de A, isto é, o conjunto $\{\mathrm{B}\colon \mathrm{B} \subseteq \mathrm{A}\}$. $\langle a, b\rangle$ denota o par ordenado cujo primeiro elemento é a e cujo segundo elemento é b. $\mathrm{A} \times \mathrm{B}$ denota o produto cartesiano dos conjuntos A e B, isto é, o conjunto $\{\langle a, b\rangle\colon a \in \mathrm{A} \text{ e } b \in \mathrm{B}\}$.

O conceito de *partição* de um conjunto é particularmente relevante neste trabalho, porque desempenha um papel importante na formulação da teoria da probabilidade. É assumido aqui conhecimento prévio desse conceito, mas cabe recordar que, se os conjuntos $\mathrm{B}_1, \ldots, \mathrm{B}_n$ são tais que para todo $a \in \mathrm{A}$ existe um único $1 \leq i \leq n$ tal que $a \in \mathrm{B}_i$, então o conjunto $\{\mathrm{B}_1, \ldots, \mathrm{B}_n\}$ é uma partição de A. Intuitivamente, pode-se imaginar que um conjunto é um bolo

[3] Recomenda-se, ao interessado nesse assunto, o livro "axiomatic set theory", de Patrick Suppes.

[4] Expressões da forma $\{a, b, c\}$ devem ser entendidas como abreviações de expressões da forma $\{\mathrm{A}\colon \mathrm{A} = a \text{ ou } \mathrm{A} = b \text{ ou } \mathrm{A} = c\}$.

e que partições são os diferentes modos de fatiar esse bolo: cada grão do bolo (elemento do conjunto) pertence a exatamente uma fatia (conjunto da partição); isto é, não existe nem um grão fora de qualquer fatia, nem um grão em mais de uma fatia. Como será adequadamente explicado no capítulo 1, as funções de probabilidade podem ser definidas tanto sobre proposições, quanto sobre conjuntos. Neste último caso, que é o mais comum em manuais dessa disciplina, deve ser possível particionar o espaço amostral; isso é necessário para a demonstração de teoremas importantes, como o teorema da probabilidade total. Se A, B são conjuntos e existe uma função injetora de domínio A e contradomínio B, então a expressão $\{a_b\}_{b \in \mathrm{B}}$ denota o conjunto A, e diz-se que B é um conjunto de *índices* de A, e que A é *indexado* por B. Por exemplo, se Γ é um conjunto enumerável, então pode ser representado pela expressão $\{g_n\}_{n \in \mathbb{N}}$.

A notação de pares ordenados é um pouco abusada para estender-se também a sequências de tamanho arbitrário, denominadas "ênuplas". Convencione-se a seguinte notação para ênuplas: se $1 \leq i \leq n$, então $\langle x_1, \ldots, x_n \rangle[i]$ denota x_i. Uma ênupla $\mathrm{D} = \langle d_1, \ldots, d_n \rangle$ está contida em uma ênupla $\mathrm{C} = \langle c_1, \ldots, c_m \rangle$ (simbolizado por $\mathrm{D} \subseteq \mathrm{C}$) se é a ênupla vazia ou se é tal que $\exists i \in \mathbb{N}^{+}(1 \leq i \leq m \wedge \mathrm{C}[i] = \mathrm{D}[1] \wedge \forall j \in \mathbb{N}^{+}(i < j \leq m \rightarrow \mathrm{C}[j] = \mathrm{D}[1 + j - i]))$. Uma ênupla D é igual a uma ênupla C (simbolizado por $\mathrm{D} = \mathrm{C}$) se $\mathrm{D} \subseteq \mathrm{C}$ e $\mathrm{C} \subseteq \mathrm{D}$. Uma ênupla D está estritamente contida em uma ênupla C (simbolizado por $\mathrm{D} \subset \mathrm{C}$) se $\mathrm{D} \subseteq \mathrm{C}$ e $\sim(\mathrm{D} = \mathrm{C})$.

Como será mais propriamente comentado no capítulo 5, a meta que norteou esta pesquisa não foi propriamente atingida aqui. Isto é, não se conseguiu demonstrar nem que há em LFI1 resultados análogos àqueles obtidos para o caso clássico, nem que não há. Verificou-se apenas que os resultados limitativos de Lewis (1976) não podem ser obtidos pela substituição da lógica proposicional clássica (LPC) pela lógica paraconsistente LFI1, já que alguns resultados técnicos não podem ser replicados apenas com trocando-se a negação clássica ($\sim$) pela negação paraconsistente ($\neg$) de LFI1. Sendo assim, o objetivo de fato atingido por esta dissertação é o de listar os resultados que continuam válidos em LFI1 e comentar quais são as dificuldades com aqueles que não podem ser reproduzidos. Para atingir esse objetivo, será feita uma análise detalhada das demonstrações indicando quais resultados não são válidos em LFI1, para que em trabalhos futuros seja possível investigar se tais resultados podem ou não ser reproduzidos nessa lógica. O intuito dessa análise é verificar se uma teoria de probabilidade baseada na lógica paraconsistentente LFI1 sofre ou não os efeitos de resultados limitativos como os demonstrados por David K. Lewis (1941–2001). Passa-se, agora, a uma descrição resumida do conteúdo e propósito de cada capítulo.

O capítulo 1 visa, sobretudo, introduzir os conhecimentos de teoria da probabilidade necessários ao entendimento da questão tematizada por esta dissertação. Tem, também, dois objetivos mais gerais: situar minimamente o leitor na história da teoria da probabilidade e do conceito de probabilidade em si, e apresentar as principais interpretações da teoria da probabilidade.

O capítulo 2 apresenta os conceitos de Lógica e Metalógica relevantes para este trabalho. A intenção é apenas relembrar ferramentas já adquiridas nessas disciplinas e fixar as convenções e notações empregadas aqui; explicações mais abrangentes e extensas podem ser buscadas em manuais de Lógica, por exemplo.

O capítulo 3 destina-se à apresentação das chamadas "leis do excesso" de Karl R. Popper (1902–1994) e dos resultados de trivialização de Lewis (1976). Estes últimos fomentaram demonstrações de resultados semelhantes. Entre eles, serão comentados também os de Fitelson (2015), que são, como argumenta o autor, os mais fortes desse tipo. As demonstrações originais de todos esses resultados são muito sucintas, uma vez que eles foram publicados como artigos. Elas serão reproduzidas aqui mais lenta e explicitamente e em português, de modo a ficarem mais acessíveis.

O capítulo 4 provê um panorama das lógicas paraconsistentes: sua história, seu lugar na Lógica e nas lógicas não-clássicas, e suas principais características formais e filosóficas.

O capítulo 5 articula os resultados do capítulo 3, que concernem a lógica proposicional clássica, e a lógica paraconsistente LFI1, a fim de investigar se aqueles teoremas são válidos ou não nessa lógica.

Finalmente, o capítulo 6 retoma as principais discussões desta dissertação, sua relevância, e elenca alguns possíveis desenvolvimentos futuros desta pesquisa.

1 Probabilidade e suas prováveis interpretações

> Per si movem-se os atomos, e logo
> Transmittem aos corpusculos mais tenues,
> E co'elles mais anologos, seu moto
> Por clandestinas impulsões causado,
> Que aos de pouco mais vulto agita e fere.
> Dos atomos assim dimana e sobe
> Gradual o movimento ate ser visto
> Por nós n'esses corpusculos, que nadam
> Do pai da luz na restea scintillante,
> Inda que as causas d'elle occultas fiquem.
>
> LUCRÉCIO, *De rerum natura*. Tradução de Antonio José de Lima Leitão (ortografia original).

Este capítulo visa apresentar os aspectos da teoria da probabilidade relevantes para o problema abordado por esta dissertação. Faz-se necessária, primeiramente, uma observação terminológica. A expressão "teoria da probabilidade" é ambígua, porque pode referir-se tanto a um determinado sistema axiomático[1], quanto a uma área da Matemática, quanto ao campo mais amplo que compreende não apenas essa área, mas também a forma com que esta é empregada na Ciência e na Tecnologia. Sendo assim, convencione-se que a expressão "Teoria da probabilidade" denotará a respectiva disciplina matemática, "teoria da probabilidade" referir-se-á às suas aplicações científicas e tecnológicas, e "teoria axiomática da probabilidade" denotará o sistema axiomático a ser apresentado na Definição 1.2.2.

Uma boa forma de começar a entender a teoria da probabilidade é compreender a sua natureza multidisciplinar. Good (1959, p. 443) argumenta que os matemáticos desenvolvem o aparato formal da teoria da probabilidade, definindo funções de probabilidade axiomaticamente; estatísticos aplicam essas funções à modelagem e solução de problemas concretos, e filósofos discutem o significado, tanto metafísico quanto epistêmico e metodológico, da definição e dos usos desse aparato. Essas áreas são independentes, mas trabalha-se melhor em cada uma tendo alguma familiaridade com as demais, argumenta o autor. Conceber a teoria da probabilidade dessa forma facilita o contato tanto com seu objeto de estudo quanto com discussões metateóricas. Ao inteirar-se de um debate dentro da teoria da probabilidade, é útil ter em mente em que região dessa tripartição da teoria ele se localiza. É mais fácil estudar os axiomas de Kolmogorov, por

[1] A teoria da probabilidade foi axiomatizada pela primeira vez por Andrey N. Kolmogorov (1903–1987) em 1933, conforme será comentado adiante.

exemplo, se se tem em mente que eles pretendem-se neutros quanto às suas possíveis aplicações na Ciência.

A teoria da probabilidade é, desde 1933, axiomática: seus axiomas prevêem como computar novos valores de probabilidade dados certos valores iniciais, mas não os estipulam. Os valores de probabilidade iniciais devem ser fixados por métodos externos à teoria da probabilidade. De fato, assim é com toda teoria axiomática: uma lógica, por exemplo, computa o valor de verdade da conjunção de duas proposições atômicas se forem conhecidos os valores de verdade individuais, mas não prevê como determiná-los. A axiomática de 1933, que deve-se a Kolmogorov, foi extremamente bem sucedida e tornou-se o cálculo de probabilidades padrão das ciências, tanto naturais quanto sociais. Pode-se dizer, por isso, que o aspecto matemático da teoria da probabilidade é praticamente incontroverso (GILLIES, 2000a, p. 1).

Já a faceta filosófica da teoria está, desde sua origem, envolta em disputas quanto a algumas grandes questões. Uma das principais é: qual a relação entre a noção informal de probabilidade e sua definição formal? Isto é, qual o significado de atribuir uma probabilidade de x% a um acontecimento? Que um agente racional ideal tem x% de confiança de que ele ocorrerá? Que um agente específico tem x% de confiança de que ele ocorrerá? Que ele ocorre em x% dos casos em que poderia ocorrer? Que suas propriedades são tais que, se as mesmas circunstâncias se repetissem muitas vezes, ele ocorreria em x% dessas vezes[2]? À ánalise das diferentes conceitualizações de probabilidade dá-se o nome de "interpretações da probabilidade": "our goal [em uma interpretação da probabilidade] is to transform inexact concepts of probability familiar to ordinary folk into exact ones suitable for philosophical and scientific theorizing" (HÁJEK, 2012). O problema da interpretação da probabilidade consiste em grande medida, portanto, em averiguar o que os valores iniciais de probabilidade significam e podem ser determinados para a teorização científica e filosófica.

Na seção 1.1, será sucintamente exposta a história da teoria da probabilidade. Na seção 1.2, será apresentado brevemente seu lado matemático, com destaque para a axiomática de Kolmogorov, de 1933. Na seção 1.3, serão discutidas algumas controvérsias da filosofia da teoria da probabilidade, com especial destaque para as interpretações da probabilidade mais populares e relevantes para esta dissertação. O terceiro aspecto da probabilidade, a sua parte mais aplicada e propriamente estatística, não será comentado aqui, por fugir ao escopo desta dissertação.

[2] Se o acontecimento em questão for o lançamento de uma moeda, as circunstâncias podem ser, por exemplo, a distribuição espacial da massa da moeda e as características físicas do dispositivo que a lança e do ambiente em que isso acontece.

1.1 História da teoria da probabilidade

> Um problema proposto a um austero jansenista por um homem do mundo foi a origem do cálculo de probabilidades.
>
> Poisson, citado por Gillies (2000a, p. 3).

A teoria da probabilidade remonta ao Iluminismo, e pode ter sua origem colocada, idealizadamente, em 1654, ano em que Blaise Pascal (1623–1662) e Pierre de Fermat (1607–1665) trocaram cartas a respeito de problemas cujas resoluções envolveram o que hoje reconhece-se como teoria da probabilidade. Restaram para a posteridade apenas três cartas de Pascal e quatro de Fermat, escritas entre julho e outubro daquele ano; não se sabe quantas cartas se perderam (DEVLIN, 2008, p. 14). Deduz-se que a correspondência começou com Pascal repassando a Fermat um problema que lhe fora apresentado por Antoine Gambaud, Cavaleiro de Méré, um matemático amador. Gambaud constatara, por participar de muitos jogos de azar, que ocorriam menos duplos seis em vinte e quatro lançamentos de dois dados do que um seis em quatro lançamentos de um dado. De fato, a probabilidade do resultado $\langle 6, 6\rangle$ em 24 lançamentos de dois dados é de apenas $1 - (\frac{35}{36})^{24} \approx 49{,}14\%$. Já a probabilidade de um seis em quatro lançamentos de um dado é $1 - (\frac{5}{6})^4 = \frac{671}{1296} \approx 51{,}78\%$[3]. Mas, como 24 está para 36 assim como quatro está para seis (isto é, $\frac{24}{36} = \frac{4}{6}$), Gambaud concluiu que há uma inconsistência na Matemática (GILLIES, 2000a, p. 5)! A resposta de Fermat perdeu-se com o tempo, mas infere-se das sete cartas recuperadas que ele resolveu corretamente o problema listando os resultados possíveis de quatro lançamentos de um dado e contando em quantos deles ocorria pelo menos um seis. Analogamente, ele deve ter listado todas as combinações possíveis de vinte e quatro lançamentos de dois dados e contado em quantas havia dois seis.

Pascal e Fermat voltaram-se, então, para um problema que ficou conhecido como "problema do jogo incompleto". A cada turno de um jogo valendo um prêmio, dois ou mais jogadores apostam, tendo chances iguais de estarem certos. A aposta pode ser prever o resultado do lançamento de uma moeda ou de um dado, por exemplo. O primeiro a vencer um determinado número de turnos vence também o jogo e embolsa o prêmio previamente estipulado. Como, em vez de contarem-se os turnos do jogo, pode-se convencionar que vencer um turno equivale a marcar um ponto, o problema do jogo incompleto é também conhecido como "problema dos pontos". Por algum motivo, o jogo não pode ser devidamente terminado: qual a forma justa de dividir o prêmio entre os participantes?

Primeiramente, Pascal e Fermat discutiram jogos de duas pessoas. Fermat observou que bastava listar todos os possíveis resultados dos turnos que não ocorreriam, e então contar em quantos cenários cada jogador ganharia. A razão entre os dois valores seria a proporção pela qual o prêmio deveria ser dividido (GILLIES, 2000a, p. 6). Por exemplo, suponha-se um

[3] O fato de que Gambaud conseguiu notar essa diferença tão sutil entre as probabilidades evidencia o quão assíduas aos jogos de azar eram algumas pessoas.

jogo que precisaria de mais três turnos para terminar, e no qual o primeiro jogador precisaria ganhar um turno para vencer, e o segundo, ganhar dois. Seja "a" um turno em que o primeiro jogador é vitorioso, e "b", um turno favorável ao segundo. Então os possíveis resultados dos três turnos restantes são: $\langle a, a, a\rangle$, $\langle a, a, b\rangle$, $\langle a, b, a\rangle$, $\langle a, b, b\rangle$, $\langle b, a, a\rangle$, $\langle b, a, b\rangle$, $\langle b, b, a\rangle$, $\langle b, b, b\rangle$. Desses oito cenários possíveis, o segundo jogador vence o jogo em apenas dois deles: $\langle b, b, a\rangle$ e $\langle b, b, b\rangle$[4]. Isso significa que ele deve ficar com $2/8 = 1/4$ do prêmio. Aqui fica claro em que consiste a relação entre o problema do jogo incompleto e o surgimento da teoria da probabilidade: dividir o prêmio corretamente exige, num certo sentido, "prever o futuro", especular sobre um término de jogo que nunca ocorrerá (DEVLIN, 2008, p. 9–10).

Além de discutirem casos de dois jogadores, Pascal e Fermat debruçaram-se sobre um jogo de três jogadores em que um deles precisa de um turno para a vitória, e os outros dois precisam, cada um, de dois turnos[5]. Pascal supõe então que o algoritmo de Fermat descrito acima, o chamado "método combinatório", é suficiente para jogos de duplas, mas não para casos em que há mais jogadores. Como sua aplicabilidade mantém-se para qualquer número de jogadores, é de se imaginar que Pascal não tinha uma resposta para o problema de Gambaud, embora tivesse dito o contrário a Fermat. Com efeito, o método combinatório leva à conclusão correta de que, nesse caso, o prêmio deve ser dividido segundo a proporção de 17 : 5 : 5 (GILLIES, 2000a, p. 6).

A origem do problema do jogo incompleto é desconhecida, mas ele foi registrado pela primeira vez em uma obra de 1494 de Luca Pacioli intitulada "Summa de arithmetica, geometria, proportioni et proportionalita" (DEVLIN, 2008, p. 8). Trata-se de um livro essencial para a Matemática e a Contabilidade renascentistas do território que hoje é a Itália. Pacioli considerou um jogo de dois jogadores que termina quando alguém vence seis turnos, mas que é interrompido quando o placar está 5 a 2. Ele concluiu que a única forma de dividir o prêmio seria distribui-lo conforme os turnos já decorridos, isto é, segundo a proporção de 5 : 2.

Mais tarde, Girolamo Cardano (1501–1576) observou corretamente que a distribuição deveria basear-se nos turnos futuros, mas também não conseguiu resolver o problema do jogo incompleto. Sua maior contribuição como precursor da teoria da probabilidade foi *Liber de ludo aleae* ("livro de jogos de azar"), escrito em 1564, mas publicado, postumamente, apenas em 1663. Embora fosse principalmente um manual para apostadores, ele já contém alguns cálculos de probabilidades.

Niccolò Tartaglia (1499/1500–1557) afirmou que o problema era insolúvel. Lorenzo Forestani propôs que parte do prêmio devia ser dividida conforme os turnos concluídos, e o restante, repartido igualmente, uma vez que o futuro não favorece ninguém em particular. A demora na aceitação das ideias de Cardano, bem como as teses de Tartaglia e Forestani, são

[4] O segundo jogador vence dois turnos também no cenário $\langle a, b, b\rangle$. Mas esse cenário não é favorável a ele porque essas vitórias jamais aconteceriam: o jogo termina assim que alguém o vence, e, nesse cenário, a vitória do primeiro jogador no primeiro dos três turnos já o faz vencedor do jogo. Os jogos devem ser formalizados como ênuplas, e não como conjuntos, justamente porque a ordem dos resultados é relevante.

[5] Embora estivessem concentrados em casos com poucos jogadores, porque para esses é fácil listar todos os cenários possíveis, a teoria de Pascal e Fermat era geral o suficiente para dar conta de qualquer número de jogadores.

evidências da força, à época, da crença de que não se pode dizer nada racionalmente significativo sobre o futuro (DEVLIN, 2008, p. 18–19).

O último precursor da teoria da probabilidade é Galileu Galilei (1564–1642). Em 1718, foi publicado um artigo seu intitulado *Sopra le scoperte dei dadi* ("sobre uma descoberta a respeito de dados"). Jogadores intuíam que, ao lançarem-se três dados de seis faces, é mais provável sua soma ser dez que ser nove. Galileu concluiu que isso é mesmo o caso, e ocorre porque há 27 modos de somarem dez, mas apenas 25 de somarem nove (GILLIES, 2000a, p. 4)[6]. Cardano já provara esse mesmo resultado, mas fizera-o de modo puramente teórico, ao invés de partir de uma observação empírica. Nesse sentido, poder-se-ia dizer que Galileu foi "mais científico" (DEVLIN, 2008, p. 19).

O motivo pelo qual essas figuras não são creditadas com a fundação da teoria da probabilidade é que seus trabalhos, ao contrário da correspondência entre Pascal e Fermat, não repercutiram de maneira duradoura, não engendraram estudos matemáticos sistemáticos (GILLIES, 2000a, p. 4). Os esforços de Pascal e Fermat fomentaram a publicação, em 1657, do primeiro tratado da História sobre teoria da probabilidade: *De raciociniis de ludo aleae* ("do raciocínio em jogos de azar"). Escrito por Christiaan Huygens (1629–1695), esse texto de dezesseis páginas foi obra de referência e introdução ao assunto durante meio século, e a partir dele a teoria da probabilidade consolidou-se como uma área da Matemática (DEVLIN, 2008, p. 98), (GILLIES, 2000a, p. 4). Muito embora ainda não a chamasse por esse nome, Huygens foi um dos primeiros a notar que a teoria da probabilidade era muito mais que uma heurística para apostas em jogos de azar: "it is not just a game which has been treated here, but [. . .] the principles and the foundations are laid of a very nice and very deep speculation" (HUYGENS *apud* (DEVLIN, 2008, p. 98)). Outro mérito de Huygens foi ter observado a importância do conceito de *expectativa* e tê-lo explicitado (DEVLIN, 2008, p. 99). Em termos informais, a expectativa de um acontecimento aleatório é a soma dos produtos entre cada um de seus resultados e o respectivo ganho. No caso do resultado ser um prejuízo, seu ganho é um número negativo. Isso fica mais claro no seguinte exemplo:

Exemplo (Expectativa de apostar R$ 100,00 nas casas vermelhas de uma roleta (DEVLIN, 2008, p. 99)). As roletas de cassinos têm 36 casas, numeradas de 1 a 36, sendo 18 vermelhas e 18 pretas. Há também duas casas verdes, marcadas com o número zero. Suponha-se que um jogador aposta R$ 100,00 nas casas vermelhas: qual a expectativa dessa aposta? Os resultados da giração da roleta relevantes para essa situação são: V (uma casa vermelha é sorteada) e ∼V (uma casa de outra cor é sorteada). O ganho associado a V é 100, já que é esse o prêmio em caso de vitória. A

[6] As possibilidades para o resultado dez são: $\langle 1,3,6\rangle$, $\langle 1,4,5\rangle$, $\langle 1,5,4\rangle$, $\langle 1,6,3\rangle$, $\langle 2,2,6\rangle$, $\langle 2,3,5\rangle$, $\langle 2,4,4\rangle$, $\langle 2,5,3\rangle$, $\langle 2,6,2\rangle$, $\langle 3,1,6\rangle$, $\langle 3,2,5\rangle$, $\langle 3,3,4\rangle$, $\langle 3,4,3\rangle$, $\langle 3,5,2\rangle$, $\langle 3,6,1\rangle$, $\langle 4,1,5\rangle$, $\langle 4,2,4\rangle$, $\langle 4,3,3\rangle$, $\langle 4,4,2\rangle$, $\langle 4,5,1\rangle$, $\langle 5,1,4\rangle$, $\langle 5,2,3\rangle$, $\langle 5,3,2\rangle$, $\langle 5,4,1\rangle$, $\langle 6,1,3\rangle$, $\langle 6,2,2\rangle$, $\langle 6,3,1\rangle$. As possibilidades para o resultado nove são: $\langle 1,2,6\rangle$, $\langle 1,3,5\rangle$, $\langle 1,4,4\rangle$, $\langle 1,5,3\rangle$, $\langle 1,6,2\rangle$, $\langle 2,1,6\rangle$, $\langle 2,2,5\rangle$, $\langle 2,3,4\rangle$, $\langle 2,4,3\rangle$, $\langle 2,5,2\rangle$, $\langle 2,6,1\rangle$, $\langle 3,1,5\rangle$, $\langle 3,2,4\rangle$, $\langle 3,3,3\rangle$, $\langle 3,4,2\rangle$, $\langle 3,5,1\rangle$, $\langle 4,1,4\rangle$, $\langle 4,2,3\rangle$, $\langle 4,3,2\rangle$, $\langle 4,4,1\rangle$, $\langle 5,1,3\rangle$, $\langle 5,2,2\rangle$, $\langle 5,3,1\rangle$, $\langle 6,1,2\rangle$, $\langle 6,2,1\rangle$.

$\sim$V associa-se -100, porque é esse o prejuízo em caso de derrota[7]. A expectativa da aposta é, portanto: $\frac{18}{38}100 + \frac{20}{38}(-100) \approx -5{,}26$. Ou seja, apostar assim implica uma expectativa de perder R$ 5,26 a cada jogo.

Um dos primeiros usos do conceito de expectativa foi o cálculo da anuidade de seguros de vida. Dado o contexto capitalista subjacente a esse tipo de contrato, um seguro de vida deve ser caro o suficiente para que a seguradora lucre, mas barato a ponto de ser desejável adquiri-lo, para o contratante. Por exemplo, uma pessoa com grandes chances de morrer em breve representa um gasto iminente à seguradora, de modo que seu seguro de vida deve ser proporcionalmente mais caro. Calcular o preço ideal requer, portanto, calcular a expectativa de vida do contratante, isto é, a idade até a qual ele provavelmente viverá. Essa probabilidade não é absoluta, mas condicional, porque depende de várias circunstâncias: sua idade atual, seu estilo de vida, o lugar em que mora, etc. As noções de probabilidade absoluta e condicional serão propriamente discutidas no capítulo 3.

Outro uso bastante antigo e bem mais interessante filosoficamente da expectativa foi a chamada “aposta de Pascal”. Trata-se de um argumento a favor não exatamente da existência de Deus, mas da crença nele (isto é, do próprio ato de acreditar) e de uma vida conduzida com base nessa crença. Pascal argumentou em seu *Pensées* que, caso o deus cristão não exista, o teísta terá desperdiçado um pouco de seu tempo e energia com uma crença errônea, mas terá como destino após sua morte o mesmo estado de inexistência que aguarda os ateus. Por outro lado, se Deus existir, então o ateu experimentará como punição uma eternidade no Inferno, ao passo que o teísta gozará de sua recompensa (ou terá potencial para tanto, ao menos). Na hipótese de Deus existir, a punição para a descrença e a recompensa para a crença são tão imensas que ultrapassam largamente quaisquer pequenos dissabores ou alegrias da hipótese contrária. Sendo assim, e como não se sabe se Deus existe ou não, é mais seguro acreditar em Deus. Isto é, a expectativa do cristianismo torna-o muito mais racional que o ateísmo. A aposta de Pascal continua a ser um argumento muito comum na Filosofia da religião e na Teologia, objeto tanto de concordância quanto de discordância. Uma crítica a ela é o fato de que a hipótese de existir uma divindade não se restringe à dicotomia entre cristianismo e ateísmo; pode ser o caso que exista um deus muito diferente do deus cristão. Assim, como escolher em qual ou quais deuses ou deusas acreditar? Outra objeção possível é que o teísmo motivado por interesse na possível recompensa pode desagradar a Deus, caso ele exista. E, se existir, o deus cristão conhece os pensamentos mais íntimos de cada pessoa e sabe se suas crenças são autênticas ou motivadas por mero interesse.

A teoria da probabilidade originou-se, como se viu, na análise de jogos de azar em meados do século XVII. Mas raciocinar sob incerteza é tão intrínseco ao pensamento que surpreende a demora para que fossem dados os primeiros passos na teoria da probabilidade,

[7] Fica claro que o cálculo de expectativa é mais simples quando os ganhos e prejuízos são monetários, porque o dinheiro é, por definição, um número. Mas a qual número associar, por exemplo, perder o braço esquerdo em um acidente de carro?

especialmente quando se considera que jogos de azar existem desde o Antigo Egito[8]. Há algumas hipóteses para o surgimento tardio da teoria da probabilidade (e para o surgimento do conceito de probabilidade em si). Duas das mais comumente aventadas serão descritas a seguir.

Demoraria até o século XIV para que cartas passassem a fazer parte dos jogos de azar, e só no começo do século seguinte começaram as primeiras análises sobre elas (quais mãos ocorrem com mais frequência, por exemplo). Muitas vezes, um matemático era contratado por um mecenas para isso (Galileu é um exemplo, conforme já comentado). Até essa época, os objetos usados nos jogos eram dados. Os dados antigos eram “astralagi” ou “tali” (no plural, “astralagus” e “talus”, no singular, respectivamente), ossos dos calcanhares de quadrúpedes que têm quatro lados planos e dois arredondados (HACKING, 2006, p. 1). Como os animais diferem um pouco entre si, a distribuição espacial da massa de seus ossos também varia um pouco, de modo que cada dado devia favorecer uma de suas faces. Isso pode ter dificultado a detecção de padrões nos resultados dos lançamentos, e, por consequência, o pensamento sistemático sobre incerteza, uma vez que os primeiros problemas da teoria da probabilidade envolviam sempre experimentos com resultados igualmente prováveis. Gillies (2000a, p. 22–23) endossa essa explicação e enfatiza que os astralagi eram irregulares, embora admita que uma minoria era entalhada de modo que as imperfeições responsáveis por viéses nos lançamentos eram removidas. Hacking (2006, p. 4), no entanto, defende que a maioria dos dados antigos era uniforme, ou porque eram astralagi lixados adequadamente, ou porque não eram feitos de ossos, mas de materiais já razoavelmente regulares, como marfim. Como se vê, não há um consenso quanto a até que ponto os jogadores antigos e medievais tinham acesso a dados justos.

Uma segunda explicação é a de que uma cosmovisão determinista coibiu o surgimento do pensamento sistemático a respeito de incerteza. O determinismo é a tese metafísica de que cada estado do Universo é tão somente produto do estado imediatamente anterior, de modo que pode ser totalmente descrito em função dele. A Ciência era entendida como um sistema de leis atemporais que relacionam matematicamente as grandezas da Natureza. O exemplo mais paradigmático desse modelo é a Física newtoniana. A lei da gravitação universal, por exemplo, fixa a relação aritmética entre a força gravitacional entre dois corpos, a distância entre eles, suas massas, e uma constante matemática. Assim, um agente que soubesse a posição e a velocidade de todas as partículas existentes e conhecesse todas as leis que as regem seria onisciente. Esse agente ideal foi imaginado pela primeira vez, ao que tudo indica, por Pierre-Simon Laplace (1749–1827), e por isso é conhecido como “demônio de Laplace”[9] (PRETO, 2015, p. 18):

8 Mas, até onde a datação dos dados rudimentares da Antiguidade permite inferir, não existiam muito antes dessa época: “we have no evidence of possible gambling devices before the Sumerian and Assyrian sites, and no proof before Egypt” (HACKING, 2006, p. 2).

9 Cabe observar que a onisciência foi, tradicionalmente, no mundo Ocidental, associada ao Deus cristão. A caracterização do agente onisciente como um “demônio” (no sentido grego antigo da palavra) é testemunha, portanto, do deísmo e do ateísmo característicos do Iluminismo e, em particular, do pensamento de Laplace. Interessantemente, a probabilidade foi usada por Pascal para argumentar em favor da crença em Deus, como se viu na discussão da chamada “aposta de Pascal”.

> *Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes (LAPLACE, 2007, p. 4).*

Dadas essa visão de mundo e concepção de Ciência, pode parecer natural que a incerteza tenha demorado tanto a encontrar lugar na Filosofia e na Ciência. Mas, para Hacking (2006, p. 2–3), um primeiro problema dessa hipótese é o anacronismo de responsabilizar uma visão de mundo do fim do Renascimento por uma ausência que data da Antiguidade. Esse anacronismo é reforçado pelo fato de que o desenvolvimento da teoria da probabilidade no século XVII é contemporâneo ao fortalecimento do determinismo alicerçado na Física da época: "far from this 'mechanical' determinism precluding an investigation of chance, it was it accompaniment" (HACKING, 2006, p. 3). Embora essa coexistência possa parecer paradoxal, fica claro que não o é quando se entende que a incerteza pode ser ou epistêmica ou metafísica.

A incerteza metafísica é aquela atribuída à realidade física "em si", que precede a observação e independe dela para existir[10]. Afirmar que existe esse tipo de incerteza é afirmar que o Universo é intrinsecamente incerto, e que, portanto, o fato de que alguns de seus fenômenos não parecem poder ser descritos senão pela teoria da probabilidade se deve, ao menos parcialmente, à própria natureza deles. Essa ideia, bastante contraintuitiva durante a maior parte de História, só foi encontrar respaldo filosófico e científico no século XX, com o desenvolvimento da Mecânica Quântica. Ao menos em algumas das interpretações dessa teoria[11], certos fenômenos são descritos em termos de probabilidade, que, então, já não representa a "chance" de um acontecimento, mas é a sua própria constituição. Tome-se como exemplo o próprio átomo, isto é, o modelo atômico. O modelo âtomico de Bohr, como o de Rutherford, que o antecedeu, assemelha-se a uma miniatura do sistema solar em que elétrons orbitam um núcleo de prótons e nêutrons. Já no modelo quântico, a posição de um elétron não-observado é não um ponto específico do espaço, mas uma função de probabilidade que traduz-se como uma "nuvem" envolvendo o núcleo. Observar o elétron ocasiona o colapso da função-onda, quando são determinadas as coordenadas de sua posição única. "Observar" não significa, nesse contexto, necessariamente presença humana, mas interação com um sistema clássico. Good (1959, p. 447) argumenta, no entanto, que a Mecânica Quântica não provou a falsidade do determinismo: "the theory of

[10] Afirmar existir incerteza metafísica pressupõe, portanto, acreditar que um mundo externo existe e pode ser conhecido diretamente, ou, pelo menos, que existe e é isomorfo a um sistema de constructos científicos que, esses sim, podem ser conhecidos diretamente. A existência de objetos externos é um problema clássico da teoria do conhecimento, e o realismo científico, por sua vez, uma discussão tradicional na Epistemologia.

[11] A Mecânica Quântica, como a Lógica e a teoria da probabilidade, é tal que certas descrições e previsões são consequência apenas do formalismo teórico, sendo, portanto, gerais, mas muito fica a cargo de interpretações. A Mecânica Quântica descrita aqui é a da interpretação de Copenhage, que é, até hoje, a opção ortodoxa, mais comumente adotada. Nem todas as suas interpretações, no entanto, afirmam a incerteza metafísica: a interpretação de de Broglie-Bohm, também conhecida como "teoria da onda piloto", é uma exceção determinista notável.

determinism is less credible than it was a hundred years ago, but is by no means disproved and never will be".

Pode-se considerar muito "caro", em termos ontológicos, uma interpretação da probabilidade fazer tantas assunções metafísicas, e preferir-se interpretações menos comprometidas com teses dessa natureza. Uma saída, nesse sentido, é justamente conceber a incerteza epistemicamente, tal que ela seja uma propriedade não de uma possível realidade física, mas do conhecimento humano. Trata-se da ideia de que as limitações que separam (e, provavelmente, sempre separarão) esse conhecimento da onisciência são o motivo pelo qual alguns fenômenos só podem ser conhecidos recorrendo-se à teoria da probabilidade. O movimento chamado browniano, por exemplo, é aquele descrito por partículas suspensas em um fluido, decorrente de colisões com outras partículas. É o tremular caótico de partículas de poeira, por exemplo, caso em que o fluido é o ar. Dada a complexidade desse tipo de movimento, a quantidade de partículas envolvidas nele, é impossível modelá-lo teoricamente tal que o padrão descrito pela partícula analisada seja dado em função do comportamento das outras partículas. Mas não é preciso, a rigor, concluir daí que a incerteza é inerente à dinâmica dos fluidos; é possível limitar-se a afirmar que o conhecimento é tal que fenômenos como o movimento browniano só podem ser conhecidos de alguma forma por meio de modelos probabilísticos.

Vê-se, assim, que uma concepção metafísica determinista não coibiria, necessariamente, o desenvolvimento da teoria da probabilidade. De fato, como será comentado na seção 1.3.2, a primeira interpretação da probabilidade, dita "clássica", contemporânea ao surgimento da teoria em si, é metafisicamente determinista: nela, é só por não sermos como o demônio de Laplace que a teoria da probabilidade tem razão de ser. Good (1959, p. 447) resume bem a relação entre a metafísica da incerteza e a epistemologia da probabilidade:

> *We can consistently talk about physical probability without committing ourselves to the metaphysical theory that the universe is indeterministic, but only if we accept the existence of subjective probability or credibility. For if we assume determinism we can get physical probabilities only by having an incompletely specified physical setup. In this incomplete specification there must be probabilities. If we are determinists we must attribute these latter probabilities to our own ignorance and not merely to something basic in nature "out there." Whether or not we assume determinism, every physical probability can be interpreted as a subjective probability or as a credibility. If we do assume determinism, then such an interpretation is forced upon us*

Determinar a causa do surgimento tardio da teoria da probabilidade é um problema em aberto da história da Ciência que escapa ao escopo desta dissertação. O objetivo dessa seção foi apenas contextualizar historicamente as discussões subsequentes. Os interessados em conhecer melhor a história da teoria da probabilidade são remetidos a Hacking (2006), sobretudo, mas também ao primeiro capítulo de Gillies (2000a) e a Devlin (2008). Como último comentário, é interessante registrar que a probabilidade pode ter sido descoberta na Índia muito antes de emergir na Europa. Hacking (2006, p. 6–8) comenta a hipótese de que no Mahābhārata, um dos dois grandes épicos indianos, é descrita uma "ciência dos dados" cujos conhecedores seriam não só favorecidos em jogos de azar, mas também capazes de fazer inferências estatísticas. Passa-se agora ao aspecto matemático da teoria da probabilidade.

1.2 Definição matemática de probabilidade

> 1. A point is that of which there is no part.
> 2. And a line is a length without breadth.
>
> EUCLIDES, *Elementos*. Tradução para o inglês de Richard Fitzpatrick.

Richard E. von Mises (1883–1953)[12] foi um precursor da visão de que a teoria da probabilidade deve ser desenvolvida independentemente de suas aplicações, isto é, ser construída como um sistema axiomático, tal qual a Geometria (GOOD, 1959, p. 445). Em teorias axiomáticas deve-se imputar a cada conceito somente aquilo que o sistema formal subjacente permite, de modo que os significados da linguagem natural não influenciem a conceitualização. Por exemplo, em Teoria de conjuntos, deve-se assumir que as palavras "conjunto" e "classe" denotam objetos cujas diferenças são exatamente as que os teoremas da teoria estipulam, ignorando-se os significados que essas palavras têm no seu uso comum[13]. Também o conceito de probabilidade não deve herdar da linguagem natural os significados usuais da palavra:

> *The theory of probability, as a mathematical discipline, can and should be developed from axioms in exactly the same way as Geometry and Algebra. This means that after we have defined the elements to be studied and their basic relations, and have stated the axioms by which these relations are to be governed, all further exposition must be based exclusively on these axioms, independent of the usual concrete meaning of these elements and their relations (KOLMOGOROV, 1950, p. 1).*

Sendo assim, a teoria da probabilidade é geralmente entendida como extensão da teoria da medida, podendo também ser definida a partir da teoria dos jogos (BUENO-SOLER; CARNIELLI, 2015, p. 16). Há dois modos de definir funções de probabilidade, mas eles são formalmente equivalentes[14]: em uma das definições, os valores de probabilidades são atribuídos

[12] Matemático e cientista austríaco que não deve ser confundido com o economista Ludwig von Mises, seu irmão mais velho.

[13] A Lógica e a teoria dos conjuntos não são exceções. De fato, a família do que poderíamos chamar "ciências formais" (Lógica, Matemática, e talvez Ciência da Computação) é apenas um caso especial, nesse aspecto como em tantos outros, da Ciência em geral. É essencial, para o pensamento científico, que algumas palavras, e cada ciência escolhe as que lhe são mais convenientes, sejam empregadas como termos técnicos; isto é, que não tenham nenhum significado senão aquele convencionado pelos cientistas. A única particularidade das ciências formais é que o significado de seus termos técnicos provém de um sistema formal, e não de experimentos ou outras fontes empíricas. Dito isso, é claro que a escolha dos termos técnicos é geralmente inspirada pela linguagem natural, até mesmo por uma questão mnemônica. Por exemplo, em Biologia, a palavra "evolução" não tem o sentido de progresso, melhoramento, que usualmente tem, mas é uma escolha conveniente porque seu significado na linguagem natural remete ao seu uso técnico.

[14] Desde que a teoria da probabilidade em questão seja clássica. Essa equivalência não necessariamente existe para teorias da probabilidade construídas a partir de lógicas não-clássicas.

a sentenças, e, no outro, são atribuídos a conjuntos (de eventos). Eles são equivalentes porque, conforme o teorema de representação de Stone, toda álgebra booleana é isomórfica a uma álgebra de conjuntos. Esse é o motivo pelo qual, ao se estudar Lógica e teoria dos conjuntos, tem-se a impressão que a disjunção "é a mesma coisa" que a união, que a conjunção equivale à intersecção, e assim por diante. As definições não são iguais, no entanto. Popper, por exemplo, interpretou a definição de probabilidade sobre conjuntos como uma subordinação da teoria às álgebras booleanas e, por isso, propôs uma nova definição (BUENO-SOLER; CARNIELLI, 2016, p. 12).

A atribuição a sentenças é a mais antiga, remontando ao interesse de Gottfried W. Leibniz (1646–1716) pelo Direito, e também a Augustus De Morgan (1806–1871), George Boole (1815–1864), entre outros (BUENO-SOLER; CARNIELLI, 2015, p. 16). Ela é, atualmente, mais comum na Lógica e na Filosofia. Tendo em vista que essas disciplinas ocupam-se de proposições, ou, pelo menos, de sentenças facilmente convertíveis em proposições, é mesmo mais natural que nelas predomine a atribuição de probabilidade a sentenças; seria desnecessariamente trabalhoso transformá-las em conjuntos.

A axiomática da probabilidade sobre conjuntos deve-se a Kolmogorov, em 1933[15], e é praticamente incontroversa (GILLIES, 2000a, p. 109). Essa definição encontra mais adeptos entre engenheiros, estatísticos, e cientistas das ciências humanas e naturais, já que os problemas a serem resolvidos com recurso à teoria da probabilidade por essas pessoas já estão, em sua maioria, configurados em termos de conjuntos (de eventos). O sistema de Kolmogorov (1950) é exibido abaixo. Antes faz-se necessário, no entanto, definir álgebras de conjuntos.

Definição 1.2.1 (Álgebra de conjuntos)**.** Se A é um conjunto, então o conjunto $\mathfrak{F}(\mathrm{A})$, denominado "álgebra de conjuntos (sobre A)", é aquele que satisfaz as seguintes condições:

i) $\mathrm{A} \in \mathfrak{F}(\mathrm{A})$;

ii) $\mathfrak{F}(\mathrm{A})$ é fechado sob complemento em relação a A e sob união finita: a diferença entre A e qualquer elemento de $\mathfrak{F}(\mathrm{A})$ e a união de quaisquer dois elementos de $\mathfrak{F}(\mathrm{A})$ também pertencem a $\mathfrak{F}(\mathrm{A})$.

Corolário 1.2.1. *Se* A *é um conjunto e* $\mathfrak{F}(\mathrm{A})$ *uma álgebra de conjuntos (Definição 1.2.1), então* $\varnothing \in \mathfrak{F}(\mathrm{A})$.

Demonstração. Devido a (i), $\mathrm{A} \in \mathfrak{F}(\mathrm{A})$. Então $(\mathrm{A} \setminus \mathrm{A}) \in \mathfrak{F}(\mathrm{A})$ (ii); mas $\mathrm{A} \setminus \mathrm{A} = \varnothing$. ■

Definição 1.2.2 (Axiomática de Kolmogorov (1950))**.** Seja Ω um conjunto enumerável, denominado "espaço amostral", cujos elementos denominam-se "eventos elementares". Seja $\mathfrak{F}(\Omega)$ uma álgebra de conjuntos sobre Ω (Definição 1.2.1) cujos elementos denominam-se "eventos aleatórios", ou simplesmente "eventos". Então seja $p\colon \mathfrak{F}(\Omega) \mapsto \mathbb{R}$ uma função, denominada "função de probabilidade", tal que:

K1 $p(\mathrm{A}) \geq 0$, para todo $\mathrm{A} \in \mathfrak{F}(\Omega)$ (não-negatividade)

[15] 1950 é o ano da tradução para o inglês consultada; o original alemão data de 1933.

K2 $p(\Omega) = 1$ (normalidade)

K3 Se $\mathrm{A}, \mathrm{B} \in \mathfrak{F}(\Omega)$ *e* $\mathrm{A} \cap \mathrm{B} = \varnothing$, então $p(\mathrm{A} \cup \mathrm{B}) = p(\mathrm{A}) + p(\mathrm{B})$ (aditividade finita)

Diz-se que $\langle \Omega, \mathfrak{F}(\Omega), p \rangle$ é um "espaço de probabilidade".

Por definição, $\mathfrak{F}(\Omega) \subseteq \wp(\Omega)$. É comum, no cálculo de probabilidades, escolher-se como álgebra de conjuntos do espaço amostral justamente o conjunto de suas partes. Essa axiomática é pensada para álgebras finitas, mas é possível trocar K3 pelo seguinte axioma:

K3' Se $\{\mathrm{A}_i\}_{i \in \mathbb{N}}$ é uma família de conjuntos dois a dois disjuntos, então $p(\bigcup_{i=1}^{\infty} \mathrm{A}_i) = \sum_{i=1}^{\infty} p(\mathrm{A}_i)$ (aditividade enumerável)

A axiomática resultante abarca álgebras de conjuntos fechadas sob união enumerável, as chamadas σ-álgebras. Diz-se que a axiomática {K1, K2, K3} é o *caso finito* da teoria de Kolmogorov, e a axiomática {K1, K2, K3'}, o *caso infinito* (PRETO, 2015, p. 14–15). A pergunta que resta é a de quais o propósito e significado dessa extensão. Kolmogorov a via como uma abstração por vezes útil, mas sem significado concreto, uma vez que todo fenômeno empírico a ser modelado pela teoria da probabilidade é naturalmente finito. Ele entende sua limitação às álgebras finitas, portanto, como uma escolha teórica arbitrária. Já Bruno de Finetti (1906–1985) admitia apenas a axiomática finita e era ativamente contrário ao caso enumerável (PRETO, 2015, p. 14–15).

Essa pergunta é parte de uma questão maior: de que maneira a teoria axiomática da probabilidade, desprovida de significado intrínseco, como todo sistema axiomático, é interpretada e aplicada à descrição de fenômenos? Isto é, quais as diferenças entre a teoria axiomática, apresentada na Definição 1.2.2, e a teoria da probabilidade como é efetivamente empregada nas ciências e engenharias? Um conceito central para a aplicação da teoria da probabilidade aos procedimentos científicos e tecnológicos é o de "experimento aleatório". Trata-se de um acontecimento de resultado incerto (dado que sujeito a forças aleatórias, ou a leis desconhecidas) que pode ser repetido indefinidamente sob as mesmas circunstâncias relevantes. Idealmente, deve ser possível identificar os parâmetros relevantes para o experimento aleatório em questão e então representá-los por variáveis aleatórias (discretas ou contínuas). No caso de uma variável discreta, são listados todos os possíveis valores que ela pode assumir, conforme o conhecimento prévio sobre a natureza do que ela representa, e cada um desses valores passa a ser um evento elementar. Uma vez determinados os conjuntos de eventos elementares correspondentes a cada variável aleatória, o espaço amostral do experimento será o produto cartesiano entre esses conjuntos. O exemplo a seguir ilustra esse procedimento:

Exemplo (Lançamento de uma moeda duas vezes (KOLMOGOROV, 1950, p. 4)). Considere-se um experimento aleatório que consiste em repetir duas vezes o lançamento de uma moeda. Convencione-se que "K" denota um resultado cara, e "C", coroa. É razoável supor que K e C são os únicos resultados possíveis de cada lançamento (não pode ocorrer de a moeda cair perfeitamente equilibrada sobre seu lado, por exemplo), de modo que o conjunto de eventos elementares é $\{K, C\}$. Então $\Omega = \{K, C\} \times \{K, C\} = \{\langle K, K\rangle, \langle K, C\rangle, \langle C, K\rangle, \langle C, C\rangle\}$ é o espaço amostral desse experimento.

A tarefa de estipular as variáveis aleatórias e os valores que cada uma assume é muito simples apenas quando o experimento em questão envolve somente fenômenos também muito simples, particularmente aptos a serem representados pela teoria da probabilidade, como o lançamento de um dado ou de uma moeda. O primeiro objeto é fabricado a fim de ser um gerador de números aleatórios, e o segundo é particularmente adequado a essa tarefa, embora não seja esse o seu intuito principal. Em experimentos mesmo que apenas ligeiramente mais complexos já é perceptível que determinar os eventos elementares exige estudo sobre metodologia experimental. Um segundo exemplo, mais próximo das aplicações concretas da teoria da probabilidade, ilustra esse fato.

Pode ser relevante para um certo estudo demográfico dividir uma população em três grupos: i) pessoas de 0 a 17 anos (grupo “jovem”), ii) de 18 a 64 anos (grupo “adulto”), e iii) de 65 anos ou mais (grupo “idoso”). Nesse caso, a idade, a princípio uma noção contínua como qualquer intervalo de tempo, foi representada por uma variável discreta que pode assumir exatamente três valores. Suponha-se ainda que há, no experimento em questão, um segundo parâmetro a ser representado também por uma variável discreta, escolaridade, e que idade e escolaridade são os únicos parâmetros desse estudo. Assumindo-se que cada pesssoa é classificada segundo o maior grau de escolaridade que obteve, os valores a serem assumidos pela variável “escolaridade” poderiam ser: i) até ensino médio (grupo I), ii) até graduação (grupo II), e iii) até pós-graduação (grupo III). Então o espaço amostral desse experimento é o conjunto $\{\text{jovem}, \text{adulto}, \text{idoso}\} \times \{\text{grupo I}, \text{grupo II}, \text{grupo III}\}$. Observe-se que nem a seleção de idade e escolaridade como variáveis aleatórias livres, nem a estipulação dos valores que cada uma delas poderia assumir, nem o cômputo do espaço amostral foram determinados pela teoria axiomática, aquela apresentada na Definição 1.2.2. Isto é, a teoria axiomática da probabilidade é silenciosa sobre como discretizar os resultados possíveis de um experimento em eventos elementares e como determinar seus valores de probabilidade. Nada é dito pela teoria, também, sobre o que significam ou representam os valores atribuídos pelas funções de probabilidade. Essas questões ficam a cargo das diferentes *interpretações* da (teoria da) probabilidade, abordadas na próxima seção: “Given certain probabilities as inputs, the axioms and theorems allow us to compute various further probabilities. However, apart from the assignment of 1 to the universal set [o espaço amostral] and 0 to the empty set, they are silent regarding the initial assignment of probabilities. For guidance with that, we need to turn to the interpretations of probability” (HÁJEK, 2012). Os protocolos de metodologia científica, no caso da ciência e tecnologia, abarcam não apenas o sistema de Kolmogorov, mas também a interpretação frequentista, a ser vista na subseção 1.3.3.

1.3 Interpretações da probabilidade

> Every axiomatic (abstract) theory admits, as is well known, of an unlimited number of concrete interpretations besides those from which it was derived.
>
> Kolmogorov (1950, p. 1).

Várias interpretações da probabilidade foram oferecidas desde o surgimento da teoria da probabilidade, descrito na seção 1.3.1. Depois de Laplace, houve quem tenha argumentado que uma das interpretações é a correta e as outras são ou enganosas, ou descrições não do conceito de probabilidade, mas de noções derivadas dele. De Finetti, por exemplo, defendeu que frequência e propensão são "pseudo-propriedades misteriosas" redutíveis à probabilidade subjetiva (HACKING, 2006, p. 14). A maioria das opiniões é mais moderada, no entanto. Os cientistas e engenheiros, que usam a teoria da probabilidade como uma ferramenta para modelar seus problemas, em sua maioria ignoram, deliberadamente ou por ignorância, as discussões envolvendo essa diversidade, e seguem os protocolos já consagrados em suas áreas de atuação sem maiores preocupações definicionais (HACKING, 2006, p. 14). No trabalho filosófico e lógico, no entanto, faz-se necessário atentar a essas questões, especialmente quando se pretende falar sobre a relação entre Lógica e teoria da probabilidade, como é o caso nesta dissertação. Nesse sentido, essa seção visa a classificar e descrever sucintamente as principais interpretações da probabilidade.

1.3.1 Taxonomia das interpretações da probabilidade

> Classification of different kinds of probability is half the problem of the philosophy of probability.
>
> Good (1959, p. 443).

A tarefa de classificação é dificultada pelo fato de que variam não só as interpretações da probabilidade, mas também os modos com que elas são distinguidas e agrupadas. Em que pese esse dissenso, podem-se delinear com segurança duas grandes famílias de interpretações, que Hacking (2006, p. 12) denomina "estatística" e "epistemológica": "on the one side it [a teoria da probabilidade] is statistical, concerning itself with stochastic laws of chance processes. On the other side it is epistemological, dedicated to assessing reasonable degrees of belief in propositions quite devoid of statistical background". Tentativas de definir um desses dois aspectos em termos do outro falharam consistentemente, de modo que essa dualidade parece inerente à própria natureza do conceito de probabilidade, seja ela qual for. Hacking compara a probabilidade a Jano, deus romano de dois rostos responsável pelo tempo e pelas transições.

As divisões internas dessas duas famílias são mais disputadas. Seria difícil opor-se, entretanto, ao destacamento, dentro da família epistemológica, da interpretação dita "clássica". Seu nome deve-se ao fato de que foi a primeira a surgir; foi contemporânea ao surgimento da própria teoria da probabilidade. Embora não tenha mais adeptos, seu estudo ainda é relevante porque contextualiza historicamente o entendimento das interpretações posteriores e da variedade de interpretações em si. Há um motivo para seu estudo ainda mais importante, entretanto. Grande parte do raciocínio cotidiano sobre incerteza se dá conforme a interpretação clássica, de modo que uma lógica que incorpore a teoria da probabilidade à sua semântica a fim de aproximar-se da racionalidade pode beneficiar-se muito do estudo da interpretação clássica. Essa interpretação será descrita na seção 1.3.2. Delimitar as demais interpretações, mais modernas, é mais difícil, até pelo fato de que no caso delas o tempo não pode servir como critério, uma vez que elas são contemporâneas umas às outras. Quais critérios poder-se-ia empregar, então? Serão apresentados alguns dos mais comumente aventados a seguir.

Dado o consenso quanto à adequação do aparato formal da teoria da probabilidade, poder-se-ia pensar que um requisito para qualquer interpretação da probabilidade é conformidade à axiomática de Kolmogorov. Essa posição foi, de fato, defendida, e ficou conhecida por "critério de admissibilidade" na literatura (PRETO, 2015, p. 16)[16]. Fazendo um paralelo com a Lógica, tal exigência é análoga à propriedade formal de *corretude* de uma lógica. Expressões como "lógica" serão precisamente definidas apenas na seção 2.1, mas adianta-se que uma lógica é correta, em sentido fraco, quando todos os seus teoremas são tautologias, isto é, quando toda proposição provada sintaticamente nela é verdadeira em todas as suas valorações possíveis. *Mutatis mutandis*, uma interpretação da probabilidade satisfaz o critério de admissibilidade se "the meanings assigned to the primitive terms in the interpretation transform the formal axioms, and consequently all the theorems, into true statements" (HÁJEK, 2012). A corretude lógica é tão incontroversa e fundamental que por vezes sequer é mencionada, ficando implícita em discussões sobre completude (que, resumidamente, é a propriedade de que toda tautologia seja um teorema). Já a admissibilidade não goza de ampla aceitação, ao contrário do que se poderia imaginar. Uma objeção, por exemplo, é a de que graus de crença racional, caso seja esse o significado atribuído à probabilidade, não são finitamente aditivos, isto é, não se comportam segundo o axioma K3 da Definição 1.2.2 (PRETO, 2015, p. 16).

Outro candidato a *desideratum* para interpretações da probabilidade é o que poderia ser denominado "determinabilidade", a saber, a existência de um método para fixar os valores de probabilidade iniciais. A vagueza da expressão "método" é proposital: o intuito é que ela seja estrita a ponto de aplicar-se somente a interpretações minimamente viáveis quanto a seus métodos, mas flexível o bastante para não barrar interpretações demais. A interpretação dita "clássica", por exemplo, estipula que, se dentre *n* acontecimentos, não há motivo para esperar ou

[16] Isto é, a expressão "critério de admissibilidade" *simpliciter* concerne o sistema de Kolmogorov, que consolidou-se como padrão científico. Mas há tantos critérios de admissibilidade quanto há axiomáticas para a teoria da probabilidade.

favorecer um em detrimento dos outros, então a probabilidade de cada um é de $^1/_n$. Isso será melhor discutido na subseção 1.3.2. Embora muito possa ser objetado contra isso, e algumas críticas serão comentadas adiante, é inegável que a interpretação clássica tem o mérito de ser, pelo menos, determinável. Nessa toada, parece também razoável exigir de uma interpretação que ela atribua um valor intermediário (isto é, maior que 0 e menor que 1) a pelo menos um evento (ou proposição): "it is essential to probability that, at least in principle, it can take *intermediate* values" (HÁJEK, 2012, grifo original). Do contrário, a teoria resultante seria nada mais que a lógica proposicional clássica, que certamente é inadequada à maioria dos fenômenos que tradicionalmente competem à teoria da probabilidade.

Há, de fato, algumas noções informais tão profundamente imbricadas com a de probabilidade, que parece necessário para qualquer interpretação da probabilidade explicitar como elas estão relacionadas. Essas ideias incluem, no mínimo, as que compõem a dualidade característica da probabilidade comentada anteriormente: *crença* (especialmente quando concebida como um contínuo de graus de crença) e *frequência*. Ou seja, uma interpretação da probabilidade precisa esclarecer porque eventos muito prováveis são também muito dignos de crença e ocorrem muito frequentemente. O mesmo raciocínio aplica-se a sentenças, caso sejam elas os objetos a que são atribuídos valores de probabilidade. Como ficará nítido nas próximas Subseções, a maior parte das interpretações define probabilidade em termos de uma dessas noções, e os teoremas obtidos a partir daí é que relacionam probabilidade às outras ideias afins. Por exemplo, na interpretação subjetiva (subseção 1.3.6), probabilidade é a quantificação da crença de um agente em uma proposição (ou na ocorrência de um evento), e a proporcionalidade entre probabilidade e frequência fica por ser explicada mais complicadamente. Já na interpretação frequentista (subseção 1.3.3), probabilidade é a frequência com que um evento ocorre em uma sequência suficientemente longa de repetições do respectivo experimento; resta, então, justificar a íntima relação entre probabilidade e (graus de) crença.

Considerados esses critérios, a Tabela 3 exibe algumas taxonomias de interpretações da probabilidade encontradas na literatura. As interpretações que aparecem mais frequentemente nas fontes consultadas são as das cinco primeiras linhas, e são elas que serão apresentadas nas próximas subseções. Isto é, não serão comentadas as interpretações distinguidas por apenas um ou dois autores.

	Hacking (2006)	Climenhaga (2019)	Good (1959)	Galavotti (2017), Preto (2015)	Hájek (2012)	Gillies (2000a)
1			clássica	clássica	clássica	clássica
2		frequentista	frequentista	frequentista	frequentista	frequentista
3		propensista	propensista	propensista	propensista	propensista
4	lógica		credibilista	lógica	lógica	lógica
5	subjetiva	subjetiva	subjetiva	subjetiva	subjetiva	subjetiva
6					*best-system*	
7						intersubjetiva
8			neoclássica			
9			graus de crença			
10			multisubjetiva			
11			tautológica			
12		graus de suporte				

Tabela 3 – As interpretações da probabilidade segundo variasárias fontes.

1.3.2 Interpretação clássica

> There is a probably apocryphal story, that when Laplace was asked by Napoleon how God fitted into this system, he replied 'Sire, I have not needed that hypothesis.' I don't think that Laplace was claiming that God didn't exist. It is just that He doesn't intervene to break the laws of Science.
>
> Stephen Hawking, em conferência intitulada "Does God play dice?".

A interpretação clássica é aquela apresentada na escola (embora geralmente não com esse nome), e possivelmente a mais intuitiva. Talvez por isso, foi a primeira a surgir, o que explica seu nome. Seu maior expoente é Laplace, que tratou do assunto em seus *Théorie analytique des probabilités* (1812) e *Essai philosophique sur les probabilités* (1814)[17] (GALAVOTTI, 2017, p. 3). Outros autores importantes foram Bernoulli, Leibniz, Pascal e Huygens. Os dois últimos são, é claro, já conhecidos.

Como já comentado, a interpretação clássica tem pressupostos deterministas, o que significa que pertence à vertente epistemológica da dualidade do conceito de probabilidade. A relação entre valores de probabilidade e conhecimento pode ser concebida de dois modos: i) pode-se entender que cada valor de probabilidade representa uma crença que um agente ideal, isto é, perfeitamente racional mas ainda sujeito às limitações inerentes à humanidade, teria; ou ii) pode-se interpretar cada valor de probabilidade como a quantificação da convicção de um indivíduo na ocorrência de um evento elementar, ou na verdade de uma proposição. Por exemplo,

[17] Laplace (2007) é uma edição recente de uma tradução para o inglês de 1901 dessa obra.

considere-se uma pessoa que está lançando uma moeda justa, mas que crê estar lançando uma moeda cuja probabilidade de resultar em cara é de 0,6. Se a probabilidade do evento elementar "cara" (ou da proposição "a moeda resultará em 'cara' ") for a crença de um agente racional perfeito, será de $^1/_2$, pois a moeda é justa. Mas se essa probabilidade for simplesmente uma representação da crença da pessoa lançando a moeda, será de 0,6, porque é assim que ela pensa (erroneamente). A interpretação clássica da probabilidade é a da primeira descrição, isto é, a crença de um agente racional perfeito. Ou seja, ainda que seja epistemológica e não estatística, é uma interpretação objetiva, porque desconsidera as opiniões individuais.

Na interpretação clássica, os valores de probabilidade iniciais (isto é, dos eventos elementares) são atribuíveis apenas sob as seguintes condições:

Definição 1.3.1 (Princípio da indiferença ou da razão insuficiente (VINEBERG, 2011, p. 713))**.** Eventos tais que não há razão para esperar ou favorecer um em detrimento do outro são considerados igualmente possíveis.

Para esses eventos, define-se probabilidade como a razão entre sua cardinalidade e a cardinalidade do espaço amostral, isto é, o número total de eventos elementares:

Definição 1.3.2 (Probabilidade segundo a interpretação clássica)**.** Para todo $\mathrm{A} \in \mathfrak{F}(\Omega)$:

$$p(\mathrm{A}) = \frac{|\mathrm{A}|}{|\Omega|}$$

Então a probabilidade de obter-se uma repetição no experimento do exemplo 1.2 é $p(\{\mathrm{KK}, \mathrm{CC}\}) = \frac{|\{\mathrm{KK,CC}\}|}{|\{\mathrm{KK,KC,CK,CC}\}|} = {}^2/_4 = {}^1/_2$. É o conhecimento das características físicas das moedas que permite-nos estipular que o espaço amostral contém apenas "cara" e "coroa", mas não "repousa sobre seu lado", por exemplo. O princípio da indiferença e a definição de probabilidade são explicitados por Laplace (2007, p. 6–7) do seguinte modo:

> *The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible.*

A interpretação clássica satisfaz o requisito de determinabilidade, portanto. Na verdade, pode ser demasiadamente determinável. Não parece muito razoável, por exemplo, assumir que sempre é possível particionar a realidade contínua em vários eventos elementares igualmente possíveis. Além disso, exigir, de cada evento elementar, que ele seja tão possível quanto os demais suscita o que é, ou parece ser, uma circularidade. A teoria da probabilidade deveria, justamente, tornar contínua a noção discreta de possibilidade; como é possível, então, que a determinação dos valores iniciais já pressuponha eventos elementares "de igual possibilidade"? Como essas possibilidades foram mensuradas, e então comparadas, se não por meio da teoria da probabilidade? Good (1959, p. 445) caracteriza a ideia de Laplace como um truque: "one of Laplace's tricks was to use the expression 'equally possible cases' instead of 'equally probable cases', and thereby to pretend that he had defined probability completely".

Outro problema da interpretação clássica é sua insensibilidade a repetições de experimento. Concluir, do fato de que a proporção de caras e coroas em muitos lançamentos de uma moeda diferiu significativamente de $1:1$, que a moeda é enviesada, parece um raciocínio correto. Mais do que isso, parece ser uma das operações básicas subjacentes ao processo de aprendizado. Uma boa interpretação da probabilidade deve validar esse tipo de inferência, portanto. Por mais vezes que uma moeda tenha sido lançada, contudo, as cardinalidades do evento "cara" e do espaço amostral continuam as mesmas, e, portanto, mantêm-se também os valores de probabilidade. Posto de outro modo, se a probabilidade de uma moeda resultar em cara é de $0{,}6 = {}^{3}\!/_{5}$, deve haver, pela interpretação clássica, $3i$ eventos elementares relevantes em um espaço amostral de cardinalidade $5i$, com $i \in \mathbb{N}^{+}$; mas que eventos seriam esses, e como determiná-los? Na interpretação clássica, a única reposta às informações fornecidas por um grande número de repetições de um experimento é a seguinte lei:

Teorema 1.3.1 (Lei da sucessão (GALAVOTTI, 2017, p. 4))**.** *Se m é o número de vezes em que um evento ocorreu e n, o número em que não ocorreu, então em um experimento repetido m + n vezes tem-se que:*

$$p(\textit{ocorrência } m+1 \mid m+n \textit{ repetições}) = \frac{m+1}{m+n+2}$$

Os problemas mais sérios para a interpretação clássica são os chamados paradoxos de Bertrand[18]. Eles não são iguais quanto ao nível de desafio que representam à interpretação clássica. De fato, se "paradoxo" for definido como uma contradição indissolúvel, como o paradoxo de Bertrand A. W. Russell (1872–1970) na teoria ingênua de conjuntos, então os paradoxos de Bertrand mais simples sequer qualificam-se como paradoxos. Aqui, no entanto, "paradoxo" será entendido à maneira de Vineberg (2011):

Definição 1.3.3 (Paradoxo e antinomia)**.** Um *paradoxo* é um argumento aparentemente válido cujas premissas são aparentemente verdadeiras, mas cuja conclusão contradiz as premissas (VINEBERG, 2011, p. 718). Uma *antinomia* é uma inconsistência em uma teoria que não pode ser desfeita senão por uma revisão dos pressupostos da teoria.

Na teoria ingênua dos conjuntos, a assunção de que a qualquer predicado lógico corresponde um conjunto implica, necessariamente, que existe um conjunto, o "conjunto de Russell", que pertence a si mesmo se e somente se não pertence a si mesmo. Na terminologia adotada aqui, o chamado "paradoxo de Russell" é não um paradoxo, mas uma antinomia, porque só pode ser evitado caso os pressupostos da teoria sejam revisados. Já os paradoxos são

[18] O nome deve-se a Joseph Bertrand (1822–1900), matemático francês.

caracterizados principalmente por explicitarem viéses na racionalidade humana, pelo menos quando exercida sem muita preocupação com rigor.

Feitas essas observações terminológicas, apresentam-se dois paradoxos de Bertrand. O primeiro é meramente uma confusão na aplicação do princípio da indiferença (Definição 1.3.1) que, uma vez desfeita, permite a atribuição correta de valor de probabilidade ao evento em questão. O segundo, no entanto, é uma possível antinomia no seio da interpretação clássica.

Problema (Problema das caixas (VINEBERG, 2011, p. 714)). Três caixas são idênticas senão pelo fato de que uma contém duas moedas de ouro, outra contém uma moeda de ouro e uma de prata, e a terceira contém duas moedas de prata. Uma caixa é escolhida aleatoriamente e uma moeda é retirada dela também aleatoriamente. Constata-se que a moeda selecionada é de prata. Então qual a probabilidade de ter sido retirada da caixa que continha duas moedas de prata? O paradoxo é que há duas respostas para essa pergunta, mas ambas parecem corretas. Convencione-se que "G" denota uma moeda de ouro, "S", uma de prata, "B", uma caixa, e que cada letra pode ter um subíndice numérico. Uma vez que a posição ou ordem das moedas dentro de cada caixa é irrelevante para o problema, as caixas podem ser representadas como conjuntos, desde que cada moeda seja identificada unicamente por um subíndice[19]. Então há três conjuntos, nomeados conforme o número de moedas de prata que contêm:

- $\mathrm{B}_0 \coloneqq \{\mathrm{G}_1, \mathrm{G}_2\}$
- $\mathrm{B}_1 \coloneqq \{\mathrm{G}_3, \mathrm{S}_1\}$
- $\mathrm{B}_2 \coloneqq \{\mathrm{S}_2, \mathrm{S}_3\}$

Resposta 1 Sendo de prata, a moeda selecionada veio ou de B_1, ou de B_2. Como não há motivo para preferir um desses eventos ao outro, o princípio da indiferença (Definição 1.3.1) dita que $p(\mathrm{B}_2) = \dfrac{|\{\mathrm{B}_2\}|}{|\{\mathrm{B}_1, \mathrm{B}_2\}|} = 1/2$.

Resposta 2 Sendo de prata, a moeda selecionada é ou S_1, ou S_2, ou S_3. O espaço amostral contém três eventos elementares, portanto, e o evento de interesse contém dois, de modo que $p(\mathrm{B}_2) = \dfrac{|\{\mathrm{S}_2, \mathrm{S}_3\}|}{|\{\mathrm{S}_1, \mathrm{S}_2, \mathrm{S}_3\}|} = 2/3$.

Qual dessas respostas é a correta?

Resolução. Formalmente, busca-se a probabilidade da caixa selecionada ser a que continha duas moedas de prata *dado que* a moeda tirada dela é de prata. Já que envolve probabilidade condicional, o problema pode ser resolvido por meio da regra de Bayes:

$$p(\mathrm{B}_2 \mid \mathrm{S}) = \frac{p(\mathrm{S} \mid \mathrm{B}_2)p(\mathrm{B}_2)}{p(\mathrm{S})}$$

É preciso determinar os três valores à direita do sinal de igualdade. As caixas têm a mesma chance de serem escolhidas, de modo que $p(\mathrm{B}_0) = p(\mathrm{B}_1) = p(\mathrm{B}_2) = 1/3$. A probabilidade de uma

[19] Outra solução seria representar as caixas como multiconjuntos e dispensar os subíndices, mas a opção adotada aqui tem a vantagem de ser muito mais simples.

moeda de prata ser tirada dado que a caixa escolhida foi a de conteúdo misto é de $p(\mathrm{S} \mid \mathrm{B}_1) = \dfrac{|\{\mathrm{S}_1\}|}{|\{\mathrm{G}_3, \mathrm{S}_1\}|} = {}^1\!/\!_2$. A probabilidade de uma moeda de prata ser tirada dado que a caixa escolhida foi a que continha duas moedas de prata é de $p(\mathrm{S} \mid \mathrm{B}_2) = \dfrac{|\{\mathrm{S}_2, \mathrm{S}_3\}|}{|\{\mathrm{S}_2, \mathrm{S}_3\}|} = 1$. Por fim, a probabilidade de uma moeda de prata ser escolhida *simpliciter* é de $p(\mathrm{S}) = \dfrac{|\{\mathrm{S}_1, \mathrm{S}_2, \mathrm{S}_3\}|}{|\{\mathrm{G}_1, \mathrm{G}_2, \mathrm{G}_3, \mathrm{S}_1, \mathrm{S}_2, \mathrm{S}_3\}|} = {}^1\!/\!_2$. Substituindo esses valores na primeira equação, obtém-se:

$$p(\mathrm{B}_2 \mid \mathrm{S}) = \frac{p(\mathrm{S} \mid \mathrm{B}_2)p(\mathrm{B}_2)}{p(\mathrm{S})} = \frac{1({}^1\!/\!_3)}{{}^1\!/\!_2} = {}^2\!/\!_3$$

E, analogamente:

$$p(\mathrm{B}_1 \mid \mathrm{S}) = \frac{p(\mathrm{S} \mid \mathrm{B}_1)p(\mathrm{B}_1)}{p(\mathrm{S})} = \frac{({}^1\!/\!_2)({}^1\!/\!_3)}{{}^1\!/\!_2} = {}^1\!/\!_3$$

Ou seja, se a moeda escolhida é de prata, então a probabilidade da caixa selecionada ser a que continha duas moedas de prata é de ${}^2\!/\!_3$. ■

O problema das caixas é, claramente, um paradoxo, e não uma antinomia, uma vez que uma segunda análise cuidadosa da questão esclarece a natureza do erro que conduz à resposta de que $p(\mathrm{B}_2) = {}^1\!/\!_2$. Esse erro consiste em não perceber que conhecer o material da moeda selecionada é obter uma informação nova, e que a probabilidade a ser calculada não é absoluta, mas condicional a essa nova informação. Ou seja, trata-se de uma aplicação indevida do princípio da indiferença (Definição 1.3.1) que, no entanto, não contesta a sua validade geral.

A familiaridade com a literatura lógico-filosófica sobre paradoxos leva à percepção de que o cerne da maioria dos paradoxos é ou autorreferência, ou o infinito. Não deve surpreender, portanto, que os principais desafios à interpretação clássica sejam os paradoxos de Bertrand que envolvem conjuntos infinitos, em particular, conjuntos incontáveis. O problema que será apresentado aqui é o da corda aleatória de um círculo.

Problema (Problema da corda aleatória (VINEBERG, 2011, p. 717)). Seja $\triangle\mathrm{ABC}$ um triângulo equilátero inscrito em um círculo centrado no ponto O. Se uma corda desse círculo for escolhida ao acaso, qual é a probabilidade de seu comprimento ser maior que o dos lados do triângulo?

Solução 1. Considere-se uma corda qualquer do círculo, e rotacione-se o triângulo de modo que o vértice B coincida com uma de suas extremidades. Então seja D sua outra extremidade. A corda $\overline{\mathrm{BD}}$ será maior que o comprimento dos lados do triângulo se e somente se D pertencer ao arco $\widehat{\mathrm{AC}}$. Como os arcos $\widehat{\mathrm{AB}}$, $\widehat{\mathrm{BC}}$ e $\widehat{\mathrm{AC}}$ formam uma partição do círculo e têm o mesmo comprimento, o princípio da indiferença (Definição 1.3.1) dita que $p(\mathrm{D} \in \widehat{\mathrm{AB}}) = p(\mathrm{D} \in \widehat{\mathrm{BC}}) = p(\mathrm{D} \in \widehat{\mathrm{AC}}) = {}^1\!/\!_3$. Então é essa a probabilidade da corda aleatoriamente selecionada ser maior que o comprimento dos lados do triângulo. Na Figura 1, $\mathrm{D} \in \widehat{\mathrm{AC}}$ e, de fato, $\overline{\mathrm{BD}}$ é maior que o comprimento dos lados do triângulo. Os mesmos passos foram seguidos para o ponto $\mathrm{D}' \notin \widehat{\mathrm{AC}}$. E $\overline{\mathrm{BD}'}$ é mesmo menor que qualquer lado do triângulo.

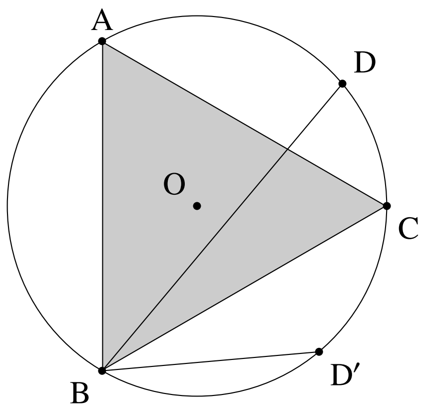

Figura 1 – A primeira resolução do problema da corda aleatória. O ponto D tem $1/3$ de chance de estar no arco $\widehat{AC}$, de modo que a probabilidade do comprimento da corda aleatória ser maior que o dos lados do triângulo é de $1/3$.

■

Solução 2. Seja D um ponto na circunferência, e rotacione-se o triângulo de modo que o raio $\overline{OD}$ fique perpendicular ao lado $\overline{AC}$ do triângulo, e seja E a intersecção entre esses dois segmentos. Então, seja F um ponto em $\overline{OD}$. Sejam G, H as intersecções entre a circunferência e a reta perpendicular a $\overline{OD}$ que passa por F. Claramente, $\overline{GH}$ é uma corda do círculo. Ela é maior que o comprimento dos lados do triângulo se, e somente se, a distância entre ela e O é menor que o comprimento de $\overline{OE}$. Como $\overline{AC}$ divide $\overline{OD}$ em segmentos de mesmo comprimento, a probabilidade de uma corda aleatoriamente escolhida ser maior que o comprimento dos lados do triângulo é de $1/2$. A Figura 2 ilustra o resultado obtido ao seguirem-se essas instruções. Observa-se que $\overline{OF}$ é menor que $\overline{OE}$, e, portanto, a corda $\overline{GH}$ é maior que o comprimento dos lados do triângulo. O mesmo procedimento foi aplicado a um outro ponto, F′, que é tal que $\overline{OF'}$ é maior que $\overline{OE}$. Como esperado, a respectiva corda $\overline{G'H'}$ é menor que o comprimento dos lados do triângulo.

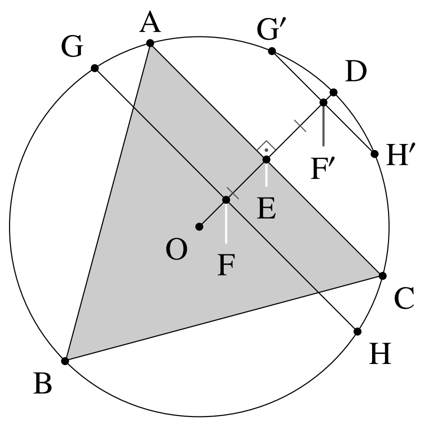

Figura 2 – A segunda resolução do problema da corda aleatória. O raio $\overline{OD}$ divide $\overline{AC}$ em segmentos de mesmo comprimento, de modo que a probabilidade do comprimento da corda aleatória ser maior que o dos lados do triângulo é de $1/2$.

■

Solução 3. Inscreva-se um círculo dentro do triângulo: ele será concêntrico ao círculo maior e seu raio será metade do raio original. Sejam D um ponto qualquer dentro do círculo e $\overline{\mathrm{EF}}$ a corda cujo ponto médio é D. Essa corda será maior que o comprimento dos lados do triângulo se, e somente se, D estiver dentro do círculo menor. Como a área do círculo menor é $^1/_4$ da do círculo maior, essa fração é também a probabilidade de que uma corda aleatória seja maior que o comprimento dos lados do triângulo. A Figura 2 ilustra um caso em que isso acontece; e, claramente, $\overline{\mathrm{EF}}$ é maior que o comprimento dos lados do triângulo. Já outro ponto qualquer, D′, está fora do círculo menor. $\overline{\mathrm{EF'}}$ é, realmente, menor que o comprimento dos lados do triângulo.

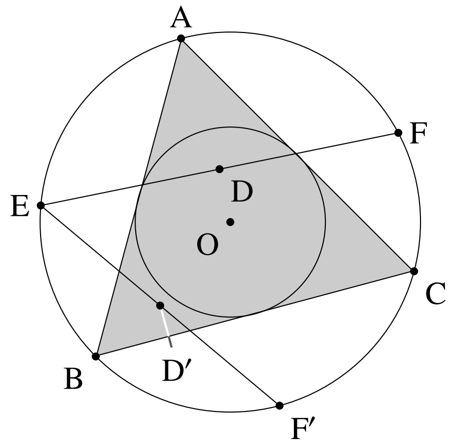

Figura 3 – A terceira resolução do problema da corda aleatória. Como a área do círculo menor é 4 vezes menor que a do maior, a probabilidade do comprimento da corda aleatória ser maior que o dos lados do triângulo é de $^1/_4$.

■

A raiz do Problema 1.3.2 parece ser o fato de que não é explicitado o modo com que a corda do círculo será escolhida. Ele deve ser aleatório, apenas, mas há pelo menos três modos de seleção que parecem igualmente aleatórios. Uma leitura atenta revela que, até nas soluções propostas, essa vagueza quanto ao método de seleção é camuflada pela ilusória boa definição de certas expressões, de modo que o problema parece solúvel à primeira vista. Na primeira resposta, por exemplo, diz-se "considere-se uma corda qualquer do círculo [. . .]"; ora, mas é crucial, para uma possível resolução definitiva do problema, saber com qual procedimento pode-se escolher uma corda do círculo. O fato de que o princípio da indiferença (Definição 1.3.1) não distingue os três algoritmos de escolha de corda apresentados parece significar que ele é falho.

1.3.3 Interpretação frequentista

> As J. M. Keynes said, "In the long run we shall all be dead."
>
> Good (1959, p. 445).

As limitações da interpretação clássica suscitaram uma busca tanto por uma teoria axiomática da probabilidade, quanto por uma concepção de probabilidade sensível à proporção entre o valor de probabilidade de um evento (ou proposição) e sua frequência em uma série de repetições de um experimento. As abordagens axiomáticas são aquelas em que o conceito central, *probabilidade*, não é definido, mas introduzido como um primitivo teórico, tal como acontece com o conceito de *conjunto* na teoria axiomática (i.e., não ingênua), de conjuntos. As conexões entre os conceitos teóricos e os fenômenos científicos são estabelecidas posteriormente, de maneira mais indireta (GILLIES, 2000b, p. 821). Já a abordagem frequentista é aquela em que probabilidade é definida em termos de frequência relativa: "in von Mises' mathematical treatment of probability, probability is explicitly defined in the mathematical formalism as limiting frequency. Kolmogorov abandons this approach, and in his mathematical development takes probability as a primitive undefined term which is characterised axiomatically" (GILLIES, 2000b, p. 821). Trata-se de uma visão *operacionalista*, portanto. O operacionalismo é a tese epistemológica de que os conceitos científicos devem ser baseados naquilo que é empiricamente observável (GILLIES, 2000b, p. 820). O frequentismo não é estritamente incompatível com uma visão axiomática, no entanto: basta que a relação de probabilidade em termos de frequência seja deslocada do formalismo matemático da teoria para a sua interpretação mais propriamente dita: "admittedly Kolmogorov's approach is still compatible with a frequency theory, if we take probability as defined in terms of frequency *in an informal supplement* designed to connect the theory with experience" (GILLIES, 2000b, p. 821, grifo próprio).

A interpretação frequentista originou-se no século XIX com os trabalhos de Robert Leslie Ellis (1843), A. Cournot (1843), Boole (1854) e John Venn (1866) (GOOD, 1959, p. 445). Nessa primeira fase, simplesmente definiu-se a probabilidade de um evento elementar como a razão entre o número de suas ocorrências em uma longa sequência de repetições do respectivo experimento e o número de repetições. Embora pareça uma boa ideia, uma falha dessa concepção é que, se uma máquina programada para lançar uma moeda gera uma sequência como $\langle K, C, K, C, K, C, K, C, K, C, K, C, K, C, K, C, K, C, K, C \rangle$ (em que "K" denota um resultado cara, e "C", um resultado coroa), nada pode ser dito senão que $p(K) = 1/2$. Qualquer observador, no entanto, questionaria imediatamente a aleatoriedade desses lançamentos (GOOD, 1959, p. 445). Problemas como esse e o prosseguimento da busca por uma teoria axiomática da probabilidade levaram à segunda fase da interpretação frequentista, já no século XX, cujo maior expoente foi von Mises (GALAVOTTI, 2017, p. 4). Ele definiu probabilidade como o limite a que tende a razão entre o número de ocorrências de um evento em um *coletivo* e a cardinalidade desse coletivo. Um *coletivo* é um objeto ideal, como o ponto na Geometria euclidiana. Mas ao contrário dele, cuja única característica é não ter dimensões, o coletivo tem uma série de propriedades, como se vê na Definição 1.3.4.

Definição 1.3.4 (Coletivo (GALAVOTTI, 2017, p. 4))**.** Um *coletivo* é uma ênupla de repetições de um experimento tal que: i) o respectivo experimento pode ser repetido indefinidamente em circunstâncias idênticas; ii) cada repetição é independente das demais; iii) a frequência do evento (ou proposição) em questão converge para um limite; e iv) o coletivo é aleatório. Uma *amostra* de um coletivo é uma ênupla estritamente contida nele. A definição de aleatoriedade de um coletivo é procedural. Um coletivo é *aleatório* se os valores de probabilidade relacionados a ele são os mesmos em qualquer uma de suas amostras: "the limiting values of the relative frequencies observed in the sub-sequences (samples) obtained by place selection equal those of the original sequence" (GALAVOTTI, 2017, p. 5).

Nota-se que a principal característica do coletivo é o fato de que cada uma de suas amostras, não importa como for selecionada, preserva os valores de probabilidades originais. Isso permite, é claro, que casos como o da máquina que produz uma sequência em que caras e coroas alternam-se perfeitamente sejam distinguidos de casos normais, ainda que em ambos a proporção de caras e coroas seja de $1:1$.

A interpretação frequentista tem seus próprios problemas, no entanto. Um dos mais óbvios é que, como os valores de probabilidade são relativos sempre a coletivos, nada pode ser dito sobre a probabilidade de experimentos que podem ser repetidos apenas uma ou algumas poucas vezes. Muitos dos fenômenos que interessam à teoria da probabilidade, entretanto, não podem ser repetidos muitas vezes sob condições idênticas. Considere-se, por exemplo, uma eleição na qual deseja-se aferir a probabilidade de um candidato ser o vencedor. Nesse sentido, o frequentismo parece estar aquém do desejável quanto ao critério de determinabilidade. Além disso, a exigência de que um experimento seja repetido "muitas vezes" é vaga. Qual o número mínimo de repetições? Good (1959, p. 445) comenta que estipular a probabilidade da giragem de uma roleta resultar em sete, dado que ela acabou de ser girada 300 vezes e em nenhuma delas ocorreu um sete, é um problema tanto para a interpretação clássica quanto para o frequentismo. A interpretação clássica é incapaz de acomodar a informação fornecida pelas 300 repetições anteriores, de modo que $p(7)$ mantém-se como $1/37$, pois é essa a razão entre o número de eventos elementares de interesse e a cardinalidade do espaço amostral. Já um frequentista pediria para a roleta ser girada mais algumas centenas de vezes antes que qualquer valor de probabilidade pudesse ser atribuído. Von Mises argumentou que a teoria da probabilidade pode ser restrita aos casos em que a exigência de grandes coletivos é factível.

1.3.4 Interpretação propensista

> We live in a world of propensities, and [. . .] this fact makes our world both more interesting and more homely than the world as seen by earlie states of the sciences.
>
> Popper, citado por Galavotti (2017, p. 7).

Popper fora um adepto da interpretação frequentista: ele defendeu essa vertente em seu "A lógica da pesquisa científica", de 1934[20]. Depois, em Popper (1957), uma comunicação em uma conferência sobre Filosofia da Física na Universidade de Bristol, ele inaugurou a interpretação dita propensista (GILLIES, 2000b, p. 807). Outro trabalho fundamental para sua nova teoria é Popper (1959) (GALAVOTTI, 2017, p. 6). Nesses trabalhos, ele argumentou que muitos fenômenos de natureza claramente probabilística não permitem repetições suficientes para que os postulados da interpretação frequentista tornem-se aplicáveis. São eventos que podem ser repetidos apenas poucas vezes, ou mesmo que ocorrem uma única vez. Nesse sentido, condicionar a atribuição de valores de probabilidade à existência de coletivos seria uma limitação severa demais à teoria da probabilidade.

Muitos fenômenos "sociais", como eleições, são exemplos de fenômenos não-repetíveis: a vitória de um dos candidatos em uma eleição é um evento que não pode ser reproduzido, porque as condições sociais que caracterizam aquela votação são uma conjuntura única. Outro exemplo é o cálculo de expectativa de vida de um indivíduo específico. Embora seja possível aferir a probabilidade de uma pessoa morrer em certa data considerando-a parte de um coletivo de pessoas que compartilham das propriedades sócio-econômicas e biológicas relevantes, falar sobre a expectativa de vida de alguém *simpliciter* não faz sentido na interpretação frequentista: "we can say nothing about the probability of death of an individual even if we know his condition of life and health in detail. The phrase 'probability of death', when it refers to a single person has no meaning at all for us" (von Mises, 1928, p. 11 *apud* Gillies (2000b, p. 809)). Não foram esses casos, contudo, a motivação de Popper, mas sim os fenômenos estudados pela Mecânica Quântica: "the main drawback of the frequency theory, according to Popper, was its failure to provide objective probabilities for single events. Yet he thought that these were needed for quantum mechanics" (GILLIES, 2000b, p. 807).

Popper desejava uma interpretação da probabilidade capaz de falar sobre eventos não-repetíveis, mas que fosse objetiva como o frequentismo, isto é, que concebesse valores de probabilidade como fatos do mundo externo, não como graus de crença racional ou relações lógicas entre proposições: "the key question for him [Popper] was whether it was possible to introduce *objective* probabilities for single events" (GILLIES, 2000b, p. 809, grifo próprio). A fim de superar a limitação da interpretação frequentista quanto a eventos não-repetíveis mantendo

[20] Esta é a data do original em alemão, *Logik der Forschung. Zur Erkenntnistheorie der modernen Naturwissenschaft*. A tradução para o inglês, publicada como *The logic of scientific discovery*, é de 1959.

a objetividade (isto é, evitando interpretações de cunho subjetivista ou lógico), o propensismo concebe a probabilidade de um evento não como sua frequência efetivamente mensurada em uma série de repetições de um experimento, mas como sua pré-disposição a ocorrer naquelas condições experimentais: "propensities may be explained as possibilities (or as measures or 'weights' of possibilities) which are endowed with tendencies or dispositions to realise themselves, and which are taken to be responsible for the statistical frequencies with which they will in fact realize themselves in long sequences of repetitions of an experiment" (POPPER, 1959, p. 30). Ou seja, a probabilidade é uma propriedade objetiva das configurações dos experimentos.

É importante não assumir que as propensões são propriedades de determinados objetos; que as propensões em um lançamento de uma moeda são características dessa moeda, ou determinadas por elas, por exemplo. A propensão de um determinado evento em um experimento é o resultado da interação de todas as circunstâncias experimentais relevantes. Por exemplo, no lançamento de uma moeda, as probabilidades de cara e coroa são efeitos não apenas das características físicas da moeda, mas também do dispositivo que a manipula e do ambiente em que isso acontece. São todas essas características do experimento que conferem àqueles lançamentos daquela moeda uma determinada propensão, ou pré-disposição, a resultar algumas vezes em cara, e outras, em coroa. Considere-se o exemplo de um dado viciado: "[...] the propensity $1/4$ *is not a property of our loaded die*. This can be seen at once if we consider that in a very weak gravitational field, the load will have little effect—the propensity of throwing a 6 may decrease from $1/4$ to very nearly $1/6$. In a strong gravitational field, the load will be more effective and the same die will exhibit a propensity of $1/3$ or $1/2$" (POPPER, 1957, p. 68, grifo original).

Um dos principais argumentos de Popper contra o frequentismo é o seguinte experimento mental (POPPER, 1959, p. 31–32): suponha-se que um dado enviesado, cuja probabilidade de resultar em seis é de $1/4$[21], foi lançado muitas vezes, mas, entre esses lançamentos, foi lançado poucas vezes (duas ou três), também, um dado justo. Dado o coletivo formado por todos esses lançamentos, qual a probabilidade de que o dado justo resulte em seis? A única resposta para a interpretação frequentista, argumenta Popper, é $1/4$, porque esse é o limite da frequência relativa da face seis no coletivo em questão (considerando-se que o número de lançamentos do dado enviesado é muito maior que o de lançamentos do dado justo). Mas essa resposta é inaceitável, porque intuitivamente sabe-se que a probabilidade do dado justo resultar em seis é a mesma em qualquer coletivo, desde que as condições experimentais não se alterem. Uma possível réplica frequentista prevista por Popper seria que a definição de coletivo deveria exigir que sequências de eventos heterogêneas como a desse caso não se qualificassem como coletivos. Isto é, um coletivo deveria ser não qualquer sequência, mas uma sequência em que cada evento resulta do mesmo conjunto de condições experimentais. Então, quando se afirma que a probabilidade do dado justo resultar em seis é de $1/4$, o que se quer dizer, na verdade, é que o limite da frequência de seis relativa a uma sequência de muitos lançamentos desse dado é

[21] Supõe-se que essa probabilidade foi aferida a partir de uma sequência anterior composta exclusivamente por muitos lançamentos desse dado.

$^1/_4$. A pergunta sobre a probabilidade do dado justo resultar em seis na sequência "mista" não faz sentido, simplesmente. Popper argumenta que essa alteração na interpretação frequentista, embora aparentemente sutil, já a torna propensista, porque desloca o conceito de probabilidade das frequências de eventos para as causas dessas frequências: "yet, if we look more closely at this apparently slight modification, then we find that it amounts to a transition from the frequency interpretation to the propensity interpretation" (POPPER, 1959, p. 34).

A interpretação propensista é melhor compreendida pela análise de seu papel no contexto da obra geral de Popper. Popper é conhecido, em ordem de importância: por seu trabalho em Filosofia da Ciência e Epistemologia, por suas ideias sobre Filosofia Política e por sua Filosofia da Física. Seu trabalho em filosofia da teoria da probabilidade é menos conhecido e originou-se de seu interesse pela Física; em particular, pelos fundamentos e pela interpretação da Mecânica Quântica. É pertinente, então, entender brevemente de que maneira a teoria propensista relaciona-se às pesquisas de Popper sobre Filosofia da Física e sobre Epistemologia.

Em linhas gerais, Popper considerava a interpretação de Copenhage da Mecânica Quântica, que naquela época consolidava-se como a interpretação mais popular, inaceitavelmente comprometida com noções deterministas e subjetivistas. Isto é, ele identificou a interpretação frequentista da probabilidade, subjacente à interpretação de Copenhage, como uma das causas de seu viés determinista e subjetivista. Nesse sentido, Popper (1959, p. 31) argumenta que a principal razão de ser do propensismo é a Mecânica Quântica, e é por sua aplicabilidade a ela que ele deve ser avaliado:

> *The main argument in favour of the propensity interpretation is to be found in its power to eliminate from quantum theory certain disturbing elements of an irrational and subjectivist character—elements which, I believe, are more 'metaphysical' than propensities and, moreover, 'metaphysical' in the bad sense of the word. It is by its success or failure in this field of application that the propensity interpretation will have to be judged. (POPPER, 1959, p. 31).*

Mais até que por sua Filosofia da Física, Popper é conhecido por sua epistemologia dita "falsificacionista". O falsificacionismo é, grosso modo, a tese de que uma afirmação é científica se, e somente se, permite estipularem-se circunstâncias que, caso ocorressem, tornariam-na falsa. Ou seja, para cada teoria científica, deve ser possível identificar uma descoberta empírica ou um resultado de experimento cuja confirmação contradiria uma previsão ou descrição da teoria, de modo a invalidá-la ("falseá-la"). Uma das características da pseudociência seria, portanto, imunidade a falseamento. A psicanálise é, para Popper, um exemplo de teoria pseudocientífica, porque suas descrições da mente humana não seriam passíveis de refutação por quaisquer experimentos ou observações empíricas. Uma teoria científica nunca é "comprovada", portanto; pode apenas, eventualmente, ser demonstrada falsa. Essa concepção contradiz a ideia possivelmente mais intuitiva, e que foi e ainda é defendida, segundo a qual as teorias científicas procedem de modo indutivo: cada nova observação que corrobora suas previsões aumenta sua probabilidade de estar correta, de modo que as teorias científicas são inferências de casos particulares para regras gerais.

Popper argumenta que o propensismo atende aos requisitos de cientificidade do falsificacionismo porque as afirmações sobre probabilidade (isto é, sobre propensão) são refutáveis por observações estatísticas. Na interpretação frequentista, frequências são a própria definição de probabilidade; no propensismo, são referências para a verificação dos valores de probabilidade: "admittedly, if we wish to *test* a probability statement, we have to test an experimental sequence. But now the probability statement is not a statement *about* this sequence: it is a statement *about* certain properties of the experimental conditions, of the experimental set-up" (POPPER, 1957, p. 68, grifo original). Ou seja, o propensismo desloca a probabilidade de frequências reais para frequências "virtuais": "these virtual frequencies [. . .] characterize the disposition, or the propensity, of the experimental arrangement to give rise to certain characteristic frequencies when the experiment is often repeated" (POPPER, 1959 apud GALAVOTTI, 2017, p. 7).

Muitas das críticas ao propensismo argumentam que distinguir frequências reais de virtuais e considerar as primeiras guias para as últimas parece implicar ou que o propensismo não é falseável, e, portanto, não é científico segundo o próprio Popper, ou que ele é redutível ao frequentismo, não tendo razão de ser. No caso de um experimento passível de muitas repetições, conceber as frequências empiricamente constatadas apenas como balizas para os valores de probabilidade não parece representar nenhum ganho teórico em relação a considerá-las diretamente os valores de probabilidade, como na abordagem frequentista. E o frequentismo tem, aqui, a vantagem de ser mais simples, devendo, portanto, ser preferido. No caso de experimentos não-reprodutíveis em larga escala, o propensismo deixa de ser científico, pelo menos segundo o falsificacionismo do próprio Popper. Isso porque não parece haver outro modo de contestar a atribuição de probabilidade de um determinado evento senão comparando-a a frequências reais, de modo que, na impossibilidade de aferir estas últimas, os valores de probabilidade tornam-se imunes a falseamento.

Poder-se-ia argumentar em defesa do propensismo que, sendo as probabilidades relativas ao desenho experimental, um método alternativo de verificar os valores de probabilidades é checar se as propriedades relevantes do experimento condizem com os valores de probabilidade hipotetizados. Mesmo que uma moeda não pudesse ser lançada, por exemplo, seria possível estudar a distribuição espacial de sua massa e, assim, avaliar a hipótese de que ela é justa (ou de que é enviesada). O problema dessa resposta é que conhecer todas as propriedades experimentais que podem influenciar os valores de probabilidade parece ser factível apenas se o experimento (no sentido técnico da palavra) em questão for, de fato, um *experimento* (no sentido geral da palavra), isto é, uma situação construída em laboratório. Novamente, ficam excluídos do escopo da teoria da probabilidade fenômenos que certamente precisam estar nele, como os resultados de eleições. Como determinar e então medir todas as variáveis relativas à chance de um certo candidato vencer uma eleição, por exemplo? Voltando ao exemplo da expectativa de vida de um indivíduo específico: a estimativa dessa expectativa deve considerar quais fatores, exatamente? O lugar em que essa pessoa vive? Sua idade atual? Seus hábitos alimentares? Seu nível de sedentarismo? Seus hábitos sociais? Enfim, como precisar todas as circunstâncias relevantes para o cálculo do valor de probabilidade? Gillies (2000b, p. 814) nomeia este problema "problema da classe de referência". A expressão "classe de referência" é mais afeita ao frequentismo, porque remete à questão análoga de como determinar a qual coletivo um determinado evento deve ter sua frequência relativa mensurada. Ainda no caso da expectativa de vida de uma certa pessoa: ao coletivo de pessoas que vivem na mesma região? De pessoas com hábitos alimentares semelhantes? E assim por diante. Esse problema evidencia, argumentam os críticos do propensismo, que valores de probabilidades não podem ser propriedades de eventos, mas de como os eventos são categorizados, o que vai contra a objetividade almejada por essa interpretação: "the problem of identifying a complete set of information is circumvented, not solved" (GALAVOTTI, 2017, p 7). Gillies (2000b, p. 815) admite que falar sobre a probabilidade de eventos não-sujeitos a repetição em larga escala, acarreta a introdução de um elemento subjetivo. Seja como for, uma das críticas mais comuns ao propensismo é a de que ele apenas adia ou desloca certos problemas inerentes à interpretação frequentista.

Como comentário final à interpretação propensista, é interessante notar que a ideia central dessa vertente já havia sido cogitada por alguns pensadores anteriores a Popper, como Charles S. Peirce (1839–1914) e o próprio Kolmogorov. Eles não são creditados pela fundação do propensismo porque falaram apenas muito superficialmente do assunto; Peirce, por exemplo, afirmou que os objetos são dotados de certos "hábitos" análogos aos das pessoas, e que os valores de probabilidades podem ser calculados a partir desses hábitos, embora não necessariamente sejam eles. Peirce incorreu no erro, portanto, de considerar as propensões como propriedades de objetos.

1.3.5 Interpretação lógica

> No general method for the solution of questions in the theory of probabilities can be established which does not explicitly recognise, not only the special numerical bases of the science, but also those universal laws of thought which are the basis of all reasoning, and which, whatever they may be as to their essence, are at least mathematical as to their form.
>
> BOOLE, *Studies in Logic and Probability.*

A interpretação lógica teve uma primeira fase, que vai de Bernard Bolzano (1781–1848) a John M. Keynes (1883–1946), passando por De Morgan, Boole, William S. Jevons (1835–1882) e William E. Johnson (1858–1931), e uma segunda fase, consolidada pela obra de Rudolf Carnap (1891–1970). Esta subseção concerne essa última etapa da história da interpretação lógica, por ser essa a sua versão mais robusta.

Nesse paradigma, a probabilidade é concebida como uma relação entre duas proposições lógicas. Ele assemelha-se à interpretação clássica por ambas as interpretações serem objetivas, ainda que epistemológicas (e não metafísicas). Isto é, nessas interpretações, as opiniões e desejos específicos dos agentes racionais não influenciam os valores de probabilidade, assim como a psicologia humana é ignorada em Lógica. Essa é a principal diferença entre as interpretações lógica e subjetiva, que será descrita na subseção 1.3.6. A probabilidade condicional é tomada, aqui, como primitiva, e é a probabilidade absoluta, unária, que é definida a partir dela. Como Carnap estava interessado sobretudo nas argumentações científica e indutiva, é comum que as proposições sejam uma *hipótese* (científica) h e uma *evidência* e. Assim, é definida uma função c tal que a proposição $c(h \mid e)$ é interpretada como o grau com que o corpo de evidências prévias e confirma[22] a hipótese h. É essa noção de grau de confirmação entre proposições que o sistema de Carnap pretende formalizar. A confirmação não necessariamente é indutiva, mas, caso esse sistema esteja correto, é um de seus méritos mais importantes a capacidade de abarcar o raciocínio indutivo. Isso porque tanto a indução quanto a incerteza são inerentes ao proceder científico e tecnológico, que são, por sua vez, sem dúvida grandes expoentes da racionalidade. Não obstante, a indução e a incerteza são de difícil tratamento lógico. David Hume (1711–1776), por exemplo, ficou conhecido por defender que não há como justificar a indução racionalmente, que seria nada mais que um hábito psicológico. Sendo um empirista radical, ele não acreditava em ideias inatas. É de se estranhar, no entanto, que ele não tenha notado que o hábito da indução, sendo ou não justificável racionalmente, é universal, o que parece favorecer a hipótese de existirem ideias inatas. O *status* da indução não melhorou linearmente com o passar do tempo. Já no século XX, Popper e David W. Miller (1942–) defenderam que a teoria da probabilidade não pode ser empregada para justificar a indução[23]. Ou seja, não procede, para eles, a visão de que a probabilidade de uma hipótese aumenta a cada acontecimento que decorre segundo suas previsões. A complexidade dos problemas relativos à (racionalidade da) indução excede os propósitos desta dissertação; essas questões são mencionadas apenas para contextualizar a interpretação lógica da probabilidade e os esforços de Carnap.

O que cabe ser observado aqui é que uma análise comum dessa formalização da noção de confirmação lógica está errada. Defende-se muitas vezes (vide Hájek (2012), por exemplo) que a lógica dedutiva é um caso particular da lógica de Carnap; mais precisamente, é o caso em que as funções de confirmação c atribuem aos seus argumentos apenas os valores 0 e 1. Ou seja, a confirmação ("*support*") $c(h \mid e)$ seria uma generalização da dedução $\{e\} \models h$ (ou da implicação material $e \rightarrow h$, considerando-se que a lógica é completa e vigora nela o metateorema da dedução: "[...] o conceito de implicação lógica é generalizado ao entendermos que $c^*(\mathrm{H}, \mathrm{E})$ determina um grau de implicação da evidência E para a hipótese H" (PRETO, 2015, p. 31). Mas isso não pode ser o caso: como será demonstrado no capítulo 3, $p(e \rightarrow h) = p(h \mid e)$ se e somente se $p(e) = 1$ ou $p(h \mid e) = 1$; nos demais casos, $p(e \rightarrow h) > p(h \mid e)$. Observado esse erro comum

[22] Não é raro ler-se que "*e* suporta *h*", mas cabe notar que "suportar" não tem (ainda?) o sentido do verbo inglês "*support*", apesar das semelhanças ortográficas e sonoras. Por esse motivo, a ideia expressa por "*support*" é representada, aqui, por palavras como "confirmar" e "embasar".

[23] Conferir Popper, Miller, "*A proof of the impossibility of inductive probability*" em *Nature*, v. 302, p. 687–688 (1983), e a polêmica suscitada por essa publicação.

que exagera as qualidades da interpretação lógica, passa-se à uma análise sucinta de seu arsenal lógico-filosófico.

As funções de probabilidade de Carnap são definidas sobre uma lógica de primeira ordem com um conjunto finito de predicados unários e um conjunto, potencialmente enumerável, de constantes de indivíduo. Cada constante c do domínio de interpretação pode ser completamente descrita por uma proposição da forma $\pm \mathrm{P}c \wedge \pm \mathrm{Q}c \wedge \pm \mathrm{R}c \ldots$, em que $\mathrm{P}, \mathrm{Q}, \mathrm{R}, \ldots$ são predicados e $\pm$ ou denota uma negação, caso o predicado em questão seja falso, ou não denota nada, caso seja verdadeiro. Essas proposições são denominadas "descrições de estado"[24]. Claramente, elas são as proposições consistentes mais expressivas da linguagem. Se $|C|, |\mathcal{P}|$ forem os números de constantes de indivíduos e de predicados, respectivamente, então haverá $2^{|C| \cdot |\mathcal{P}|}$ descrições de estados (a base 2 deve-se à bivalência da lógica clássica). Cada proposição da linguagem é equivalente a exatamente uma forma normal disjuntiva, ou seja, a uma proposição que consiste de disjunções de várias descrições de estado. Por isso, as funções de probabilidade, definidas originalmente sobre descrições de estado, são estendíveis a todas as proposições da linguagem (HÁJEK, 2012). Seguir-se-á parcialmente, aqui, a notação de Carnap, de modo que m denotará uma função de probabilidade. Para cada m, define-se a respectiva função de confirmação c:

$$c(h \mid e) \coloneqq \frac{m(h \wedge e)}{m(e)}$$

c é simplesmente a probabilidade condicional padrão do cálculo de Kolmogorov, baseada em conjuntos, adaptada à semântica de proposições. Há infinitas funções m e, consequentemente, infinitas funções c. Carnap argumenta que uma dessas m, no entanto, se sobressai por sua conformidade com a intuição a respeito do raciocínio indutivo e probabilístico. Essa função é denominada m^*, e a respectiva função de confirmação é denominada c^*. Convencione-se que uma *descrição de estrutura* é um conjunto maximal de descrições de estado. Assim, se houver

[24] Mais rigorosamente falando, as descrições de estado não são exatamente as proposições compostas apenas por conjunções de predicados negados ou não, mas as classes de equivalência dessas proposições. Ou seja, dadas a comutatividade e a associatividade da conjunção, é desejável que as proposições $(\mathrm{P}a \wedge \neg \mathrm{P}b) \wedge \mathrm{P}c$, $(\neg \mathrm{P}b \wedge \mathrm{P}a) \wedge \mathrm{P}c$, $\mathrm{P}a \wedge (\neg \mathrm{P}b \wedge \mathrm{P}c)$, $\neg \mathrm{P}b \wedge (\mathrm{P}a \wedge \mathrm{P}c)$, etc., sejam representações diferentes da mesma descrição de estado. Por simplicidade, elas continuarão a ser descritas como proposições, mas deve-se manter em mente que são, a rigor, classes de equivalência.

apenas um predicado (P) e exatamente três constantes de indivíduo (a, b, c), então as $2^{1\cdot 3} = 8$ descrições de estado serão:

1. $\mathrm{P}a, \mathrm{P}b, \mathrm{P}c$
2. $\neg\mathrm{P}a, \mathrm{P}b, \mathrm{P}c$
3. $\mathrm{P}a, \neg\mathrm{P}b, \mathrm{P}c$
4. $\mathrm{P}a, \mathrm{P}b, \neg\mathrm{P}c$
5. $\mathrm{P}a, \neg\mathrm{P}b, \neg\mathrm{P}c$
6. $\neg\mathrm{P}a, \mathrm{P}b, \neg\mathrm{P}c$
7. $\neg\mathrm{P}a, \neg\mathrm{P}b, \mathrm{P}c$
8. $\neg\mathrm{P}a, \neg\mathrm{P}b, \neg\mathrm{P}c$

E as descrições de estrutura serão:

I. $\{1\}$
II. $\{2, 3, 4\}$
III. $\{5, 6, 7\}$
IV. $\{8\}$

m^* atribui um mesmo peso a cada descrição de estrutura, e esse peso é, por sua vez, dividido igualmente entre cada descrição de estado pertencente à respectiva descrição de estrutura.

	Descrição de estado	Descrição de estrutura	Peso	m^*
1	$\mathrm{P}a \wedge \mathrm{P}b \wedge \mathrm{P}c$	I	$1/4$	$1 \cdot 1/4 = 1/4$
2	$\neg\mathrm{P}a \wedge \mathrm{P}b \wedge \mathrm{P}c$			$1/3 \cdot 1/4 = 1/12$
3	$\mathrm{P}a \wedge \neg\mathrm{P}b \wedge \mathrm{P}c$	II	$1/4$	$1/3 \cdot 1/4 = 1/12$
4	$\mathrm{P}a \wedge \mathrm{P}b \wedge \neg\mathrm{P}c$			$1/3 \cdot 1/4 = 1/12$
5	$\mathrm{P}a \wedge \neg\mathrm{P}b \wedge \neg\mathrm{P}c$			$1/3 \cdot 1/4 = 1/12$
6	$\neg\mathrm{P}a \wedge \mathrm{P}b \wedge \neg\mathrm{P}c$	III	$1/4$	$1/3 \cdot 1/4 = 1/12$
7	$\neg\mathrm{P}a \wedge \neg\mathrm{P}b \wedge \mathrm{P}c$			$1/3 \cdot 1/4 = 1/12$
8	$\neg\mathrm{P}a \wedge \neg\mathrm{P}b \wedge \neg\mathrm{P}c$	IV	$1/4$	$1 \cdot 1/4 = 1/4$

Tabela 4 – Os valores de m^* e os pesos das descrições de estados e de estrutura para uma linguagem com apenas um predicado (unário) e exatamente três constantes de indivíduo. Adaptada de Hájek (2012).

Note-se que m^* define uma c^* que se atualiza com a experiência, porque $c^*(\mathrm{P}c \mid \mathrm{P}a) = 2/3$, ao passo que $c^*(\mathrm{P}c) = 1/2$.

1.3.6 Interpretação subjetiva

Há uma íntima relação, seja qual for a sua natureza, entre probabilidade, frequência relativa e graus de crença (de um agente). O frequentismo é a identificação entre os dois primeiros conceitos; analogamente, a interpretação subjetiva, também chamada personalismo e Bayesianismo subjetivo (HÁJEK, 2012), é a definição de probabilidade em termos de graus de crença. Como as crenças variam não só entre agentes mas também, para cada agente, ao longo do tempo, as respectivas funções de probabilidade precisam ser indexadas de acordo.

Uma abordagem, denominada por Hájek (2012) "subjetivismo irrestrito", é a de representar as crenças diretamente, isto é, sem impor qualquer restrição de coerência, adequação aos axiomas de Kolmogorov, etc. A Psicologia tem mostrado que as intuições a respeito de probabilidade tendem a desviar-se muito das previsões do cálculo de probabilidades, de modo que a adoção do subjetivismo irrestrito acarretaria, certamente, em valores de probabilidade incompatíveis com regras como a aditividade finita (axioma K3 da Definição 1.2.2). A viabilidade, a utilidade e o significado de construir um sistema formal nesses termos são questionáveis. Hájek (2012) limita-se a dizer que o subjetivismo irrestrito não é "uma proposta séria", mas essa não é uma questão tão simples: qualquer formalização do comportamento humano que pretendesse modelá-lo e prevê-lo corretamente precisaria ter uma representação precisa de suas bases credenciais, mesmo que apenas para compará-las a um sistema axiomático e então descartá-las como irracionais. Seja como for, é verdade que, mesmo que o subjetivismo irrestrito possa ser justificado, não há motivo para considerá-lo um sistema de probabilidade (GOOD, 1959, p. 446). *Mutatis mutandis*, um modelo psicológico das intuições sobre condicionais contrafactuais não necessariamente qualifica-se como uma lógica, por mais útil e preciso que possa ser.

A versão do subjetivismo interessante à teoria da probabilidade, que poder-se-ia chamar "subjetivismo restrito", é aquela em que os graus de crenças são limitados por critérios de racionalidade. Nesse sentido, a teoria da probabilidade seria incorporada à Lógica como a "lógica da crença parcial" (HÁJEK, 2012). É preciso, então, estabelecer o que vêm a ser "(graus de) crenças". Na Psicologia, o movimento *behaviourista* (que foi hegemônico durante décadas) defendia que estados mentais, como emoções, não podem ser objeto de estudo, por não serem observáveis. Os *comportamentos* dos indivíduos é que deveriam ser a unidade mínima de análise, porque é mediante eles que esses indivíduos respondem às circunstâncias que os acometem. Guardadas as devidas proporções, a definição subjetivista de "(graus de) crença" lembra o paradigma *behaviourista* porque é dada em termos não de estados mentais, mas de comportamento:

> *Let us suppose that an individual is obliged to evaluate the rate p at which he would be ready to exchange the possession of an arbitrary sum S (positive or negative) dependent on the occurrence of a given event E, for the possession of the sum* pS*; we will say by definition that this number p is the measure of the degree of probability attributed by the individual considered to the event E, or, more simply, that p is the probability of E (according to the individual considered; this specification can be implicit if there is no ambiguity) (DE FINETTI* apud *(HÁJEK, 2012)).*

Claramente, o evento E pode ser convertido em uma proposição *e* que é verdadeira se e somente se o evento E ocorre. A fidedignidade de *p* como representação da crença, em um determinado momento, de um agente em *e* depende, é claro, desse agente ignorar o valor de S, de modo que não tenha outro critério para associar-lhe um valor senão sua intuição (PRETO, 2015, p. 25). Assim, a atribuição de p pode ser esquematizada como na Figura 4:

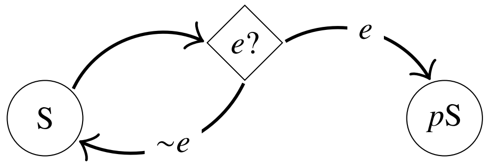

Figura 4 – A atribuição de uma probabilidade p a uma proposição e.

2 Lógica e lógicas

> Now reasoning is an argument in which, certain things being laid down, something other than these necessarily comes about through them.
>
> ARISTÓTELES, *Tópicos*. Tradução de W. A. Pickard-Cambridge.

Este capítulo apresenta, por um lado, os principais conceitos de Lógica relevantes para a compreensão das discussões desta dissertação tanto sobre lógica clássica quanto sobre lógicas não-clássicas. Por outro lado, introduz mais especificamente aspectos sintáticos e semânticos de LPC, que são importantes porque os resultados de excesso e trivialização analisados no capítulo 3 tratam de uma LPC equipada com uma semântica probabilística. O objetivo geral e o específico são atingidos simultaneamente.

2.1 Sintaxe de lógicas proposicionais

> A leitura dos clássicos, que não falam de poentes, tem-me tornado inteligíveis muitos poentes, em todas as suas cores [. . .]. No fundo é a mesma coisa—a capacidade de distinguir e de subtilizar. Sem sintaxe não há emoção duradoura. A imortalidade é uma função dos gramáticos.
>
> FERNANDO PESSOA, *Livro do desassossego*.

A definição de uma teoria formal axiomática está condicionada à existência de um conjunto de símbolos que, concatenados conforme certas regras gramaticais, compõem expressões da teoria. É preciso ainda haver um conjunto de certas expressões, denominadas *axiomas* dessa teoria, tais que exista um método computável para decidir se uma dada expressão é ou não um axioma. Por fim, deve haver ainda um conjunto finito não-vazio de relações entre expressões. Cada relação é denominada uma *regra de inferência*, e diz-se que sua última expressão é *consequência direta* da(s) anterior(es) (MENDELSON, 2015, p. 27).

Cabe ressaltar que os símbolos ainda não têm nenhum significado, ou seja, existem de modo meramente sintático. O conjunto de símbolos pode ser particionado em três: um conjunto de símbolos de variáveis proposicionais, um conjunto de símbolos de conectivos (denominado *assinatura*), e um conjunto de símbolos auxiliares (ou extra-lógicos). Este último conjunto contém simplesmente símbolos como parênteses, colchetes, e vírgulas, que facilitam e organizam a leitura das expressões formais, mas são dispensáveis. Convencione-se o seguinte conjunto de símbolos auxiliares: {"(", ")", "[", "]", ","}. A seguir, serão definidos os outros dois conjuntos; depois, serão propriamente definidos os conceitos de axioma e regra de inferência.

Seja $\mathcal{P} := \{p_n\}_{n\in\mathbb{N}}$ um conjunto enumerável de símbolos denominados *variáveis proposicionais*, que serão denotadas pelas letras $p, q, r, \ldots$, com ou sem subíndices numéricos. Assim, temos, por exemplo: p_1, q_1, r_2, p_2, etc. Como as variáveis proposicionais não têm qualquer propriedade intrínseca a não ser a de serem diferentes umas das outras[1], o mesmo conjunto $\mathcal{P}$ será compartilhado por todas as lógicas que serão introduzidas (assim como o conjunto de símbolos auxiliares, é claro).

Seja Σ uma família $\Sigma = \{\Sigma_\mathrm{n}\}_{n\in\mathbb{N}^+}$. Para todos $\Sigma_\mathrm{m}, \Sigma_\mathrm{n} \in \Sigma$, $\Sigma_\mathrm{m} \cap \mathcal{P} = \varnothing = \Sigma_\mathrm{n} \cap \mathcal{P}$ e, se $m \neq n$, então $\Sigma_\mathrm{m} \cap \Sigma_\mathrm{n} = \varnothing$. Denomina-se Σ uma *assinatura*. Os elementos de cada Σ_n são denominados *conectivos n-ários*, ou *conectivos de aridade n*. Todo conectivo tem aridade maior ou igual a 1[2]. Pode-se, enfim, definir as regras de formação que determinam quais concatenações de símbolos são expressões bem-formadas, e quais não são.

Definição 2.1.1 (Linguagem (proposicional) (adaptada de Mendelson (2015, p. 28)))**.** Sejam $\mathcal{P}$ um conjunto de variáveis proposicionais e Σ uma assinatura. Uma *linguagem* $\mathcal{L}$ é a álgebra construída por Σ sobre $\mathcal{P}$ de acordo com as seguintes regras:

i) se $\alpha \in \mathcal{P}$, então $\alpha \in \mathcal{L}$;
ii) para todo $c \in \Sigma_\mathrm{n}$, se $\{\alpha_1, \ldots, \alpha_n\} \subseteq \mathcal{L}$, então $c(\alpha_1, \ldots, \alpha_n) \in \mathcal{L}$;
iii) os elementos de $\mathcal{L}$ são exatamente aqueles gerados pelas duas regras anteriores.

Cada objeto $\alpha \in \mathcal{L}$ é denominado uma fórmula bem formada (fbf) de $\mathcal{L}$, ou uma $\mathcal{L}$-fórmula. Uma fbf também pode ser denominada uma "proposição" ou "sentença". As fbfs geradas pela primeira regra são denominadas "proposições (ou sentenças) *atômicas*". As demais são chamadas "proposições (ou sentenças) *moleculares*".

Como foi convencionado que todas as linguagens proposicionais compartilham o mesmo conjunto $\mathcal{P}$ de variáveis proposicionais (e os mesmos símbolos auxiliares), o único modo de duas linguagens diferirem é terem assinaturas diferentes. Posto de outro modo, duas linguagens (proposicionais) são iguais se, e somente se, são geradas pela mesma assinatura. Proposições arbitrárias serão representadas por letras gregas minúsculas ($\alpha, \beta, \gamma, \ldots$). Trata-se de uma convenção para a metalinguagem em que os conceitos formais (como variáveis proposicionais, assinatura e linguagem) estão sendo definidos. Como de praxe, empregar-se-á notação infixa para conectivos binários. Ou seja, proposições da forma $c(\alpha, \beta)$, em que c é um conectivo binário, serão escritas como $(\alpha\ c\ \beta)$. Quando não houver risco de ambiguidade,

[1] Isto é, o único sentido inerente a expressões como "p_1" e "p_2" é serem concatenações diferentes de símbolos (nesse exemplo, devido aos seus últimos caracteres). Todo "significado" atribuído a elas é externo, oriundo das regras sintáticas e/ou semânticas que regem seu funcionamento.

[2] É comum, em manuais de Lógica, haver um conjunto Σ_0 cujos elementos são denominados "constantes". Essa convenção não é empregada aqui por dois motivos. Em primeiro lugar, considera-se que introduzir como "conectivos 0-ários" objetos que serão interpretados como constantes torna desnecessariamente confusas tanto a linguagem-objeto sendo definida, quanto a metalinguagem em que isso é feito. Em segundo lugar, essa convenção é mais afeita a linguagens de primeira ordem (ou de ordem superior).

os parênteses correspondentes a conectivos unários e binários serão omitidos, de modo que escrever-se-á, por exemplo, $\sim\alpha$ e $\alpha\, c\, \beta$, em vez de $(\sim\alpha)$ e $(\alpha\, c\, \beta)$, respectivamente.

Um modo equivalente de construir uma linguagem é estipular um *alfabeto* de símbolos, um conjunto cujos elementos são, no mínimo, as variáveis proposicionais em $\mathcal{P}$ e os conectivos em Σ, mas possivelmente também os símbolos auxiliares. Então uma fórmula (não necessariamente bem formada) é uma sequência finita de símbolos desse alfabeto. Dado o conjunto de todas as fórmulas, a respectiva linguagem é o seu menor subconjunto que contém todas as variáveis proposicionais e é fechado pelos conectivos da assinatura.

Resta introduzir um tipo de conjunto crucial, a relação de dedutibilidade, para então passar-se à definição de "lógica".

Definição 2.1.2 (Relação de dedutibilidade)**.** Sejam $\mathcal{L}$ uma linguagem, A, B dois de seus subconjuntos e α, β duas de suas proposições. Então $* \subseteq \wp(\mathcal{L}) \times \mathcal{L}$ é uma *relação de dedutibilidade* para $\mathcal{L}$ caso satisfaça as seguintes cláusulas:

(reflexividade) se $\alpha \in A$, então $A * \alpha$[3];

(monotonicidade) se $A * \alpha$, então $(A \cup B) * \alpha$;

(corte) se $A * \alpha$ e $(B \cup \{\alpha\}) * \beta$, então $(A \cup B) * \beta$.

Comentário. Quando estiver claro a qual lógica o símbolo $\vdash$ refere-se, expressões da forma $\alpha \dashv\vdash \beta$ abreviam expressões da forma "$\{\alpha\} \vdash \{\beta\}$ e $\{\beta\} \vdash \{\alpha\}$". *Mutatis mutandis* para expressões da forma $\alpha \;=\!\!|\!\!\models \beta$. Além disso, expressões da forma "$\Gamma \not\models \alpha$" abreviam "não é o caso que $\Gamma \models \alpha$".

Definição 2.1.3 (Lógica, teoria)**.** Uma *lógica* L é uma terna $L = \langle \mathcal{L}, \vdash, \models \rangle$, em que $\mathcal{L}$ é uma linguagem construída como na Definição 2.1.1 e $\vdash, \models$ são relações de dedutibilidade definidas conforme a Definição 2.1.2.
Uma *teoria* de uma lógica L, também denominada "L-teoria", é um subconjunto de sua linguagem. Uma teoria é *estrita* se está estritamente contida na respectiva linguagem.

Ocasionalmente chama-se de "lógica" a sua linguagem, por metonímia: nesse uso menos rigoroso, pode-se dizer, por exemplo, que uma teoria Γ é subconjunto de uma lógica L (embora rigorosamente seja subconjunto da linguagem de L). Seja $\mathcal{L}$ uma linguagem, $\Gamma \subseteq \mathcal{L}$ uma de suas teorias e $\alpha \in \mathcal{L}$ uma de suas proposições. Se $\Gamma \vdash \alpha$, diz-se que Γ *deriva* α, ou que α é *teorema* de Γ. Analogamente, se $\Gamma \models \alpha$, diz-se que Γ *modela* α, ou que α é *consequência* de Γ. Uma proposição é um *axioma* de uma lógica se é teorema de sua teoria vazia. Uma teoria é *fechada* se contém todas as suas consequências, isto é, se $\Gamma \vdash \alpha$ implica $\alpha \in \Gamma$. Uma lógica é fechada se todas as suas teorias estritas são fechadas.

[3] Utiliza-se notação infixa também no nível metalinguístico, para relações binárias. Isto é, se R é uma relação binária, escreve-se $x \mathrel{R} y$ em vez de $\langle x, y \rangle \in R$.

Diz-se que uma lógica L de linguagem $\mathcal{L}$ é *correta*, em sentido fraco, se, para todas fbfs $\alpha, \beta \in \mathcal{L}$, $\alpha \vdash_{L} \beta$ implica $\alpha \models_{L} \beta$. Ou seja, quando todos os seus teoremas são tautologias. Analogamente, diz-que uma lógica L de linguagem $\mathcal{L}$ é *completa* se, para todas fbfs $\alpha, \beta \in \mathcal{L}$, $\alpha \models_{L} \beta$ implica $\alpha \vdash_{L} \beta$. Isso significa que todas as suas tautologias podem ser demonstradas axiomaticamente.

Definidos esses conceitos gerais, introduzem-se agora uma assinatura específica e a linguagem correspondente, a linguagem de LPC. Dado $\mathcal{P}$, LPC pode ser construída a partir de várias assinaturas, de modo que a escolha feita aqui é um pouco arbitrária. Convencione-se que a expressão "lógica clássica" referir-se-á não só a LPC, mas também à lógica clássica de primeira ordem (também conhecida como cálculo de predicados).

A assinatura adotada para LPC será o conjunto $\Sigma_{\text{LPC}} = \{\sim, \rightarrow\}$. A linguagem de LPC será o conjunto gerado por Σ_{LPC} sobre $\mathcal{P}$, e será denotada por $\mathcal{L}_{\text{LPC}}$. $\mathcal{L}_{\text{LPC}}$ contém fbfs como $\sim(p \rightarrow \sim p)$, $\sim\sim p$, $p_1 \rightarrow p_2$, mas elas ainda não têm nenhum significado. O modo com que serão interpretadas será explicado na próxima seção.

Definição 2.1.4 (Axiomática de LPC (MENDELSON, 2015, p. 28–29))**.**

i) $\alpha \rightarrow (\beta \rightarrow \alpha)$
ii) $(\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow ((\alpha \rightarrow \beta) \rightarrow (\alpha \rightarrow \gamma))$
iii) $(\sim\beta \rightarrow \sim\alpha) \rightarrow ((\sim\beta \rightarrow \alpha) \rightarrow \beta)$

Regra de inferência: $\{\alpha \rightarrow \beta,\ \ \alpha\} \models_{\text{LPC}} \beta$ (*modus ponens*)

Outros conectivos usuais e relevantes para esta dissertação são definidos a partir da negação e da implicação, de modo que cada ocorrência de um desses conectivos não-primitivos deve ser entendida como abreviação de uma fbf em que ocorrem apenas negações e implicações (MENDELSON, 2015, p. 29). As intuições semânticas por trás de todos os conectivos serão propriamente discutidas na seção 2.2; por ora, eles têm uma existência meramente sintática.

Conjunção $\alpha \wedge \beta := \sim(\alpha \rightarrow \sim\beta)$

Disjunção $\alpha \vee \beta := \sim\alpha \rightarrow \beta$

Equivalência/bi-implicação $\alpha \leftrightarrow \beta := (\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)$

Uma vez fixadas uma assinatura (e, portanto, uma linguagem) e uma axiomática para LPC, passa-se à análise dos teoremas de LPC necessários para a demonstração dos resultados dissecados no capítulo 3. Esses teoremas clássicos serão enunciados e avaliar-se-á em que medida as propriedades que LPC não compartilha com LFI1 são necessárias para as demonstrações dos resultados de excesso e trivialização tematizados neste trabalho, e se essas eventuais dependências poderiam ser simuladas ou contornadas dentro de LFI1. Em particular, esses teoremas não serão demonstrados; recomenda-se, ao interessado nessas provas, Mendelson (2015) e o apêndice de Bueno-Soler (2004).

Os primeiros resultados são bastantes simples, e afirmam a comutatividade, a associatividade e a idempotência da conjunção. Pode-se pensar intuitivamente que a comutatividade significa que a ordem em que ocorrem os argumentos de um conectivo binário (como a conjunção) é irrelevante para efeito de demonstrações, porque a fbf em uma das ordens deriva a respectiva fbf na outra. Já a associatividade mostra que, dados duas ocorrências adjacentes de um mesmo conectivo binário, é indiferente dar precedência à resolução de um ou de outro. Ou seja, a maneira com que "os parênteses são organizados" não importa, contanto que os argumentos da expressão maior não sejam alterados. Por fim, é idempotente todo conectivo unário c que, ao tomar como argumento uma expressão α cujos conectivos são todos c, forma uma expressão equivalente a α (isto é, intersubstituível em relação a α). Para estender essa noção a conectivos binários, basta dizer que um conectivo binário é idempotente se sucessivas composições de sua respectiva relação consigo mesma resultam sempre na relação original. Formalmente, temos:

Teorema 2.1.1 (Comutatividade da conjunção)**.** $\alpha \wedge \beta \leftrightarrow \beta \wedge \alpha$.

Teorema 2.1.2 (Associatividade da conjunção)**.** $\alpha \wedge (\beta \wedge \gamma) \leftrightarrow (\alpha \wedge \beta) \wedge \gamma$.

Teorema 2.1.3 (Idempotência da conjunção)**.** $(\alpha \wedge \alpha) \leftrightarrow \alpha$.

Essas propriedades da conjunção valem não só em LPC, mas, de fato, na maioria das lógicas, principalmente, é claro, naquelas mais semelhantes ao caso clássico. LFI1 é uma lógica muito próxima de LPC (num sentido que será especificado mais adiante), e esses teoremas concernem exclusivamente a conjunção, de modo que é de se esperar que a conjunção de LFI1 também seja comutativa, associativa e idempotente.

Teorema 2.1.4 (Conjunção tautológica)**.** $\alpha \wedge \top \leftrightarrow \alpha$.

Teorema 2.1.5 (Implicação em termos da negação e da disjunção)**.** $(\alpha \rightarrow \beta) \leftrightarrow ({\sim}\alpha \vee \beta)$.

Como será visto na Definição 3.0.1, a axiomática adotada para implementar a interface entre LPC e teoria da probabilidade probabilística privilegia conectivos como disjunção e conjunção em detrimento da implicação material. Por isso, as implicações frequentemente são introduzidas como expressões em que ocorrem apenas negações e disjunções.

Teorema 2.1.6 (Leis de De Morgan)**.**

Disjunção $\alpha \vee \beta \leftrightarrow {\sim}({\sim}\alpha \wedge {\sim}\beta)$

Conjunção $\alpha \wedge \beta \leftrightarrow {\sim}({\sim}\alpha \vee {\sim}\beta)$

As chamadas "leis de De Morgan", em homenagem ao lógico homônimo, são teoremas de LPC que especificam como a disjunção pode ser expressa em termos da negação e da conjunção, e como a conjunção pode ser expressa em termos da negação e da disjunção. Ou seja, essas duas regras explicitam que há uma dualidade entre esses dois conectivos clássicos. Como já comentado, os teoremas do capítulo 3 lidam majoritariamente com negações, disjunções e conjunções, de modo que muitas vezes torna-se necessário “converter” um desses conectivos binários no outro.

Teorema 2.1.7 (*Modus ponens* internalizado)**.** $[(\alpha \rightarrow \beta) \wedge \alpha] \leftrightarrow \alpha \wedge \beta$.

Esse teorema é interessante porque *modus ponens* é a única regra de inferência da axiomática de LPC adotada aqui, e esse teorema pode ser encarado como uma interiorização dessa regra dentro da linguagem-objeto, dado o metateorema da dedução clássico[4]. Isto é, $[(\alpha \rightarrow \beta) \wedge \alpha] \leftrightarrow \alpha \wedge \beta$ equivale, pelo metateorema da dedução, ao argumento $\{\alpha \rightarrow \beta, \alpha\} \models_{\overline{LPC}} \alpha \wedge \beta$, que é outra maneira de enunciar a regra de *modus ponens*.

Teorema 2.1.8 (Princípio do terceiro excluído)**.** $\alpha \vee {\sim}\alpha$.

Esse teorema é a expressão em LPC de uma regra que existe desde a lógica grega antiga. Naquela época, foi considerada um princípio básico da racionalidade humana, e era um dos três princípios da lógica aristotélica, ao lado do princípio da identidade e do princípio da não-contradição. De maneira talvez contraintuitiva, esse teorema é válido também em LFI1, assim como em outras lógicas paraconsistentes e não-clássicas, como será comentado adiante.

Teorema 2.1.9 (Expansão)**.** $\alpha \leftrightarrow (\alpha \wedge \beta) \vee (\alpha \wedge {\sim}\beta)$.

No método de dedução natural, a regra de introdução da disjunção permite obter-se uma fórmula como $\alpha \vee \beta$ a partir de α ou de β, e a regra de introdução da conjunção deriva $\alpha \wedge \beta$ a partir de α e β. Esse teorema assemelha-se um pouco a essas duas regras, na medida em que introduz uma disjunção de conjunções a partir de uma fórmula atômica. Ele assemelha-se muito ao princípio do terceiro excluído.

Teorema 2.1.10 (Princípio da não-contradição (simples))**.** ${\sim}(\alpha \wedge {\sim}\alpha)$.

Assim como o já mencionado princípio do terceiro excluído, esse teorema é a expressão em LPC de um dos três princípios da lógica aristotélica, ao lado do princípio da identidade e do princípio do terceiro excluído. Mas seu papel era particularmente proeminente, e, a despeito do surgimento das lógicas paraconsistentes, ainda é parte do senso comum a ideia de que contradições impossibilitam qualquer pensamento racional[5]. Um dos axiomas da Definição 3.0.1 determina que, caso a conjunção de duas fbfs seja uma partícula *bottom*, então a probabilidade da disjunção delas é a soma de cada probabilidade. Dado o princípio de não-contradição, que vigora em LPC, conjunções da forma $\alpha \wedge {\sim}\alpha$ sempre são partículas *bottom*, de modo que pode-se igualar a probabilidade de proposições como $\alpha \vee {\sim}\alpha$ à soma das probabilidades de α e ${\sim}\alpha$, e essa

[4] O metateorema da dedução é o resultado metalógico, dado que está acima da linguagem-objeto clássica, segundo o qual se Γ é um conjunto de fbfs e α e β são fbfs tais que $\Gamma \cup \{\alpha\} \models_{\overline{LPC}} \beta$, então $\Gamma \models_{\overline{LPC}} \alpha \rightarrow \beta$. Em particular, se $\{\alpha\} \models_{\overline{LPC}} \beta$, então $\models_{\overline{LPC}} \alpha \rightarrow \beta$ (MENDELSON, 2015, p. 30).

[5] Para um exame pormenorizado da história do princípio de não-contradição, recomenda-se Gomes (2013).

igualdade é necessária em várias passagens das demonstrações daquele capítulo. Dada a própria natureza e motivação das lógicas paraconsistentes, é a dependência de certas demonstrações à contraditoriedade de conjunções como $\alpha \wedge {\sim}\alpha$ que espera-se ser o maior empecilho à transposição dos resultados do capítulo 3 para LFI1.

Teorema 2.1.11 (Princípio da não-contradição (complexo))**.** ${\sim}((\alpha \wedge \beta) \wedge (\alpha \wedge {\sim}\beta))$.

Teorema 2.1.12 (Eliminação da dupla negação)**.** ${\sim}{\sim}\alpha \leftrightarrow \alpha$.

O Teorema 2.1.11 é simplesmente o resultado de algumas operações de comutação e associação sobre o princípio da não-contradição simples (Teorema 2.1.10. Já o Teorema 2.1.12 é a explicitação do fato de que, em LPC, negar uma proposição um número ímpar de vezes equivale a negá-la uma única vez, e negá-la um número par de vezes equivale a afirmá-la (isto é, não negá-la). Eliminar ou introduzir duplas negações é uma correção necessária em várias passagens de uma demonstração na qual empregou-se uma regra que altera o número de negações de uma proposição (as leis de De Morgan, por exemplo).

2.2 Semântica de lógicas proposicionais

> The claim of a formal system to be a logic depends, I think, upon its having an interpretation according to which it can be seen as aspiring to embody canons of valid argument [. . .].
>
> Haack (1978, p. 3).

A única diferença, até aqui, entre conectivos e variáveis proposicionais é que os primeiros são símbolos necessariamente distintos dos últimos e estão, cada um, associados a um número, que é sua aridade. É preciso, no entanto, que as proposições sejam classificadas quanto a sua verdade, isto é, que sejam associadas a *valores de verdade*. Essa associação será obtida pela definição de funções denominadas “valorações”.

Definição 2.2.1 (Valoração)**.** Sejam Σ uma assinatura e Δ um conjunto não-vazio de símbolos denominados “valores de verdade” tais que $\mathcal{P} \cap \Delta = \varnothing = \Sigma \cap \Delta$. Então uma *valoração v* é uma função $v \colon \mathcal{P} \mapsto \Delta$.

Assim, são atribuídos significados a todas as proposições atômicas (embora esses “significados” sejam, apenas, a associação a um valor de verdade). As lógicas não estipulam quais valores de verdade suas proposições atômicas recebem, porque ignoram os conteúdos dessas proposições, dizem respeito apenas às propriedades formais delas. De fato, uma proposição atômica sequer tem um conteúdo, no sentido cotidiano da palavra.

O próximo passo, que é valorar (isto é, atribuir um valor de verdade) proposições moleculares, pode ser feito de várias formas. Em lógicas modais munidas de uma semântica de mundos possíveis de Kripke, por exemplo, a determinação do significado de uma proposição modal em um mundo W depende do valor de verdade da respectiva proposição nos mundos acessíveis a W. A semântica apresentada aqui é bem mais simples, no entanto, em boa parte por ser *verofuncional*. Uma valoração é dita verofuncional se é uma função, isto é, se o valor de verdade de cada proposição molecular é determinado pelo(s) valor(es) de verdade da(s) variável(is) proposicional(is) que ocorre(m) nessa proposição.

Para cada valoração v construída conforme a Definição 2.2.1, será definida uma extensão v' que, por abuso de notação, depois será denominada simplesmente v. Assim estendidas, as valorações abrangirão também as proposições moleculares. É preciso, primeiramente, que cada conectivo de uma assinatura seja associado a uma *função de verdade*.

Definição 2.2.2 (Interpretação)**.** Sejam Σ uma assinatura e Δ um conjunto não-vazio de valores de verdade definido como na Definição 2.2.1. Então seja Ξ um conjunto tal que, para todo $c \in \Sigma$, se n é a aridade de c, então existe $\xi\colon \Delta^n \mapsto \Delta$ em Ξ. Então uma *interpretação* (da assinatura Σ) é uma função[6] $\Sigma \mapsto \Xi$, e os elementos de Ξ são denominados "funções de verdade".

Definição 2.2.3 (Valoração (estendida))**.** Sejam Σ uma assinatura e $\mathcal{L}$ a linguagem gerada por Σ sobre $\mathcal{P}$. Sejam Δ um conjunto não-vazio de valores de verdade construído conforme a Definição 2.2.1, $i\colon \Sigma \mapsto \Xi$ uma interpretação (para algum conjunto Ξ de funções de verdade), e $v\colon \mathcal{P} \mapsto \Delta$ uma valoração. Então defina-se a função $v'\colon \mathcal{L} \mapsto \Delta$:

$$v'(\alpha) := \begin{cases} v(\alpha) & \text{se } \alpha \text{ é uma fbf atômica} \\ i(c)(v'(\beta_1), \ldots, v'(\beta_n)) & \text{se } \alpha \text{ é uma fbf da forma } c(\beta_1, \ldots, \beta_n) \\ & \text{(e } c \text{ é um conectivo n-ário)} \end{cases}$$

Agora é possível equipar semanticamente $\mathcal{L}_{\text{LPC}}$, a linguagem da lógica proposicional clássica.

Definição 2.2.4 (Semântica de LPC)**.** Sejam $\mathbf{2} := \{\mathtt{t}, \mathtt{f}\}$ um conjunto de valores de verdade e $v_{\text{LPC}}\colon \mathcal{P} \mapsto \mathbf{2}$ uma valoração. Seja $\Xi_{\text{LPC}} := \left\{\overline{\sim}\colon \mathbf{2} \mapsto \mathbf{2},\ \curlyvee\colon \mathbf{2}^2 \mapsto \mathbf{2}\right\}$ um conjunto de funções de verdade tais que:

$$\overline{\sim}(\star) := \begin{cases} \mathtt{t} & \text{se } \star = \mathtt{f} \\ \mathtt{f} & \text{se } \star = \mathtt{t} \end{cases} \qquad \curlyvee(\star, *) := \begin{cases} \mathtt{t} & \text{se } \star = \mathtt{t} \text{ ou } * = \mathtt{t} \\ \mathtt{f} & \text{do contrário} \end{cases}$$

[6] Embora as interpretações bijetoras sejam um pouco mais intuitivas (afinal, qual seria a utilidade de definir vários conectivos com a mesma função de verdade?), a bijeção não é necessária.

Então $i_{\mathrm{LPC}}\colon \Sigma_{\mathrm{LPC}} \mapsto \Xi_{\mathrm{LPC}}$ é uma interpretação para $\mathcal{L}_{\mathrm{LPC}}$ tal que $i_{\mathrm{LPC}} = \{\langle \sim, \overline{\sim} \rangle, \ \langle \vee, \curlyvee \rangle\}$. Assim, a respectiva $v'_{\mathrm{LPC}}\colon \mathcal{L}_{\mathrm{LPC}} \mapsto \mathbf{2}$ é definida como:

$$v'_{\mathrm{LPC}}(\alpha) := \begin{cases} v_{\mathrm{LPC}}(\alpha) & \text{se } \alpha \text{ é uma fbf atômica} \\ \overline{\sim}(v'(\beta)) & \text{se } \alpha \text{ é da forma } {\sim}\beta \\ \curlyvee(v'(\beta), v'(\gamma)) & \text{se } \alpha \text{ é da forma } \beta \vee \gamma \end{cases}$$

Qual o intuito das definições de $\mathbf{2}$, $\overline{\sim}$ e $\curlyvee$? Com $\mathbf{2}$, objetiva-se representar "a verdade" e "a falsidade". $\overline{\sim}$ formaliza a ideia expressa por uma frase muito comum em textos de Lógica, Filosofia, Matemática e áreas afins: "não é o caso (que alguma coisa)". Muitos usos da palavra "não", em português, também são corretamente apreendidos pela negação de LPC: a oração "não vou sair de casa agora" é, de fato, interpretada como sinônima de "não é o caso que eu vou sair de casa agora". Não é preciso ir muito longe, no entanto, para se deparar com construções como "não vou sair de casa agora, *não*!", que, não obstante a dupla negação, tem as mesmas condições de verdade de "não vou sair de casa agora". E o português não é uma exceção: as línguas naturais não são facilmente convertíveis em proposições lógicas, e por isso existe a Semântica Formal, uma área da Linguística dedicada, entre outras coisas, a investigar as possibilidades de fazê-lo. A intuição por trás de $\curlyvee$ é o que se chama "disjunção inclusiva", uma cláusula com duas condições que é satisfeita caso pelo menos uma delas seja satisfeita. Palavras como "ou" denotam disjunções inclusivas, em certos casos; em outros, esse papel é desempenhado por "e", por exemplo. Compare-se "à noite a gente pode ir ao cinema e jantar" a "estou planejando sair pra almoçar ou ver um filme amanhã". Em particular, $\curlyvee$ *não* expressa a ideia de alternativas mutuamente excludentes, como na oração "você pode ficar com seu celular novo ou obter seu dinheiro de volta".

Alguns outros conectivos são bastante consagrados na lógica clássica, e podem ser definidos a partir de Σ_{LPC}:

Definição 2.2.5 (Σ_{LPC}). Seja $\Sigma_{2'} = \{\wedge, \rightarrow, \leftrightarrow\}$ um conjunto de conectivos interpretados por $\Xi_{\mathrm{LPC'}} = \left\{\left\langle \wedge, \curlywedge\colon \mathbf{2}^2 \mapsto \mathbf{2} \right\rangle, \ \left\langle \rightarrow, \overline{\rightarrow}\colon \mathbf{2}^2 \mapsto \mathbf{2} \right\rangle, \ \left\langle \leftrightarrow, \overline{\leftrightarrow}\colon \mathbf{2}^2 \mapsto \mathbf{2} \right\rangle\right\}$ tais que:

$\star$	$*$	$\curlywedge(\star, *)$
t	t	t
t	f	f
f	t	f
f	f	f

$\star$	$*$	$\overline{\rightarrow}(\star, *)$
t	t	t
t	f	f
f	t	t
f	f	t

$\star$	$*$	$\overline{\leftrightarrow}(\star, *)$
t	t	t
t	f	f
f	t	f
f	f	t

Então sejam $\Sigma_{\mathrm{LPC'}}$ a assinatura $\Sigma_{\mathrm{LPC}} \cup \Sigma_{2'}$ e $\mathcal{L}_{\mathrm{LPC'}}$ a linguagem gerada por $\Sigma_{\mathrm{LPC'}}$ sobre $\mathcal{P}$.

De fato, esses conectivos são tão comuns, que, a partir de agora, a lógica LPC não será mais mencionada, e a expressão "LPC" denotará a lógica LPC′. Dá-se continuidade, então, à análise dos significados dos conectivos de LPC.

O sentido de $\wedge$ é bastante intuitivo: trata-se de uma cláusula com duas condições que é satisfeita se, e somente se, ambas o são. Já a leitura de $\overline{\rightarrow}$ é mais problemática. De fato, não há exagero em dizer que boa parte das controvérsias envolvendo a lógica clássica se deve ao significado da implicação material. A intuição subjacente a proposições da forma $\alpha \rightarrow \beta$ é a de que β (chamado *consequente* da implicação) é o caso sempre que α (o *antecedente* da implicação) o for. Suponha-se que a proposição b representa a ideia expressa pela oração "a rua X está molhada" e que a simboliza o conteúdo proposicional de "está chovendo na região da rua X". Então parece muito razoável afirmar que $a \rightarrow b$. Isso significa que, sempre que está chovendo na região da rua X, a rua X está molhada. Note-se que nada é dito sobre se está ou não chovendo na rua X no momento em que $a \rightarrow b$ é enunciada. Pelos mesmos motivos, pode soar necessário admitir que, se chover na região da rua X e, ao mesmo tempo, a rua estiver seca, então será falso que $a \rightarrow b$. Esses são os dois casos mais "amigáveis" da implicação material: $\overline{\rightarrow}(\mathtt{t},\mathtt{t}) = \mathtt{t}$ e $\overline{\rightarrow}(\mathtt{t},\mathtt{f}) = \mathtt{f}$. O que acontece quando é falso que esteja chovendo na região da rua X (isto é, $\sim a$)? Como nos dois casos anteriores, há apenas duas alternativas, mutuamente excludentes: ou $v(a \rightarrow b) = \mathtt{t}$, ou $v(a \rightarrow b) = \mathtt{f}$[7]. A segunda opção parece inaceitável: não faz sentido rejeitar, devido a um dia sem chuva, que sempre que chove na região da rua X, então a rua X fica molhada. As implicações nada dizem sobre os casos em que seus antecedentes são falsos, de modo que não é razoável considerá-las falsas nessas circunstâncias. É forçoso admitir, portanto, que, se $v(a) = \mathtt{f}$, então $v(a \rightarrow b) = \mathtt{t}$ e, em geral, se $v(\alpha) = \mathtt{f}$, então $v(\alpha \rightarrow \beta) = \mathtt{t}$. Observe-se, ainda, que, do contrário, a implicação "se o imperador do Japão é paulista, então ele é brasileiro", por exemplo, seria falseada pelo fato de que o imperador do Japão não é paulista. Isso é bastante indesejável, porque soa correto que, se ele fosse paulista, então seria (também) brasileiro (trata-se de um condicional contrafactual).

Em que pesem esses argumentos a favor da razoabilidade da implicação material, algumas consequências de sua semântica são bastante contraintuitivas—e, consequentemente, controversas. Note-se, por exemplo, que quaisquer duas proposições verdadeiras implicam uma à outra, de modo que "se $2 + 2 = 4$, então Curitiba é a capital do Paraná" expressa uma proposição verdadeira. Intuitivamente esperar-se-ia, no entanto, que os antecedentes "tivessem algo a ver" com os respectivos consequentes, de modo que implicações "esdrúxulas" como essa ou fossem consideradas falsas ou, talvez melhor que isso, não recebessem um valor de verdade. Na lógica clássica, no entanto, toda proposição tem um valor de verdade (e há exatamente dois valores de verdade, de modo que proposições da forma $\alpha \vee \sim\alpha$ são tautológicas). A expectativa de relevância também é contrariada pelo fato de que cada contradição implica uma proposição qualquer[8].

[7] Há um duplo abuso de notação, aqui. Em primeiro lugar, escreveu-se v em vez de v_{LPC}. Além disso, omitiu-se o acento que distingue v, uma função cujo domínio é $\mathcal{P}$, de v', uma função cujo domínio é $\mathcal{L}_{\mathrm{LPC}}$. Quando usadas devidamente, essas abreviações mais auxiliam a leitura e o entendimento que dificultam, de modo que, quando não houver risco de ambiguidade, subíndices serão omitidos e assumir-se-á que toda valoração é estendida (isto é, construída conforme a Definição 2.2.3).

[8] Trata-se de um ponto fundamental para esta dissertação. Esse fato é a razão pela qual inconsistência e trivialidade coincidem em sistemas clássicos. O capítulo 4 detalhará essa equivalência, mas desde já pode-se perceber porque isso acontece.

O que essa discussão sobre a compatibilidade entre a implicação material e o raciocínio humano evidencia é que a semântica desse conectivo é perfeita para a formalização de argumentos lógicos, matemáticos ou afeitos a essas disciplinas, mas incongruente com argumentos de outros âmbitos da racionalidade, em especial o pensamento sobre contrafactualidade e causalidade. É exatamente essa, na verdade, a motivação para muitas das lógicas não-clássicas, inclusive as lógicas probabilísticas abordadas nesta dissertação. A próxima seção detalha a relação entre lógica e pensamento sob incerteza.

Definição 2.2.6 (Demonstração (MENDELSON, 2015, p. 27))**.** Em uma lógica L, uma *demonstração* de uma fbf α é uma ênupla de fbfs $\langle\beta_1, \ldots, \beta_k\rangle$ tais que $\alpha = \beta_k$ e, para todo $1 \leq i \leq k$, ou β_i é um axioma de L, ou é consequência direta das proposições $\beta_1, \ldots, \beta_j$, para $1 \leq j < i$, dadas as regras de inferência de L.

Definição 2.2.7 (Teoria trivial)**.** Uma teoria Γ é trivial se toda fórmula bem formada α é seu teorema: Γ é trivial $\iff \forall\alpha(\Gamma \models \alpha)$.

Definição 2.2.8 (Teoria consistente)**.** Uma teoria Γ é consistente se não há fórmula bem formada α que seja seu teorema e cuja negação também o seja: Γ é consistente $\iff \nexists\alpha(\Gamma \models \alpha \text{ e } \Gamma \models {\sim}\alpha)$.

Uma lógica é chamada trivial se todas as suas teorias o são; o mesmo vale para a consistência. É finitamente trivializável se tem teorias triviais. Isso significa que toda lógica explosiva é finitamente trivializável; lógicas não-explosivas podem ser finitamente trivializáveis ou não.

Uma proposição é uma partícula *bottom* se torna trivial uma lógica quando é adicionada a ela. Isto é, dada uma lógica L e uma proposição α, α é uma partícula *bottom* se $L \cup \{\alpha\} \models_{L} \beta$ para toda β. Qualquer partícula *bottom* pode ser denotada por um único símbolo, convencione-se $\perp$, sem risco de ambiguidade, porque todas as partículas *bottom* são equivalentes entre si. Se uma partícula *bottom* é axioma de uma lógica, essa lógica é trivial.

Uma proposição é uma partícula *top* se é consequência de toda teoria, inclusive da teoria vazia. Será denotada, quando existir, por $\top$. Novamente, não há risco de ambiguidade. Claramente, todos os axiomas de uma lógica são partículas *top* suas.

2.3 Lógica clássica e incerteza

> Such standard formalizations of probability admittedly fail to resolve important questions about uncertainty and belief, as much as standard logic alone admittedly fails to resolve all questions about truth and truth preservation. So, new forms of logic, probability or their combinations are justifiable.
>
> Walter Carnielli, *Tendencies in Logic, and some modest advice to young logicians*. Felsefe Arkivi—Archives of Philosophy, v. 51, n. 343–350, 2019.

Como comentado na seção 1.2, os sistemas axiomáticos lidam, em princípio, apenas com os objetos cujas existências são estipuladas por seus axiomas, e esses objetos são manipulados conforme os respectivos teoremas (e somente conforme eles). LPC, por exemplo, é um sistema contendo alguns esquemas de axiomas e tendo, geralmente, *modus ponens* como única regra de inferência[9]. Mas evidentemente há, por trás desses sistemas sem significado intrínseco, frequentemente a intenção de que eles modelem formalmente algum fenômeno empírico. Isto é, excetuando-se os sistemas cuja motivação exclusiva é uma exploração formal metateórica, manipulam-se os axiomas e as regras de inferência de modo que os teoremas obtidos formalizem o mais precisamente possível o fenômeno desejado. A teoria da probabilidade, por exemplo, almeja axiomatizar as intuições por trás do conceito de "chance (de algo acontecer)", mas as interpretações divergem muito quanto à natureza exata desse conceito, conforme discutido no capítulo 1. Já em lógicas, os axiomas modelam os conectivos da assinatura que, por sua vez, têm (pretensamente) a capacidade de formalizar noções extra-lógicas. Mais especificamente, as lógicas são frequentemente construídas a fim de representarem o funcionamento de argumentos, raciocínios estruturados segundo a noção de "o que se segue do quê" (o que se chamaria de *entailment*, em inglês):

> *It is relevant to distinguish, at the outset, between* interpreted *and* uninterpreted *formal systems: uninterpreted, a formal system is just a collection of marks, and cannot, therefore, be identified as a formal logic rather than, say, a formalisation of a mathematical or physical theory. The claim of a formal system to be a logic depends, I think, upon its having an interpretation according to which it can be seen as aspiring to embody canons of valid argument [. . .] (HAACK, 1978, p. 3, grifos da autora).*

Se a Lógica ocupa-se primeiramente da formalização da ideia de acarretamento, qual a explicação para a força, mesmo atualmente, da ideia de que trata exclusivamente das inferências de cunho matemático? Trata-se de um fenômeno muito mais histórico que propriamente lógico. A Lógica contemporânea foi fundada em meados do século XIX por Boole, Peirce e F. L.

[9] Várias outras axiomatizações são possíveis, no entanto, inclusive uma em que há apenas um axioma e várias regras de inferência.

Gottlob Frege (1848–1925) sob a égide do *logicismo*, um projeto que pretendia demonstrar que a Matemática é, em última análise, Lógica; isto é, que pode ser representada por uma lógica consistente. O êxito dessa corrente garantiria que a Matemática forma um todo coeso, e que suas diferentes disciplinas não divergem entre si (que as conclusões da Geometria não contradizem as da Análise, por exemplo). Soa muito plausível que essas metas sejam alcançáveis; o que seriam equações, por exemplo, senão proposições afirmando igualdades?

Os esforços nesse sentido foram abalados, no entanto, no começo do século XX, pela descoberta de paradoxos nos fundamentos da Matemática, em particular, na teoria dos conjuntos (que era, até então, a teoria ingênua de Georg F. L. P. Cantor (1845–1918)) (MARCOS, 2010, p. 138). Esses contratempos iniciais não chegaram a destruir o logicismo, mas fomentaram o surgimento de dois programas epistemológicos alternativos. Um era o *intuicionismo* de Luitzen E. J. Brouwer (1881–1966). Trata-se, grosso modo, da tese de que o raciocínio matemático precisa ser construtivo; que provar α exige construir uma prova para α, tal que provas por absurdo são inválidas. Fundou-se, assim, uma nova lógica, com axiomas e regras de inferência pensados para que seus teoremas fossem a contraparte lógica do raciocínio matemático (também intuicionista). A outra alternativa foi o formalismo de David Hilbert (1862–1943), que visava a, primeiramente, demonstrar a consistência da aritmética (MARCOS, 2010, p. 136). Talvez o marco mais importante dessas buscas tenha sido o *Principia Mathematica*, a imensa obra de Alfred N. Whitehead (1861–1947) e Russell que visava a nada menos que completar o projeto logicista ao construir uma lógica cujos teoremas fossem exatamente as verdades matemáticas. Tudo isso provou-se fadado ao fracasso quando Kurt F. Gödel (1906–1978) demonstrou seus dois teoremas da incompletude, que deixam claro que um sistema forte o suficiente para expressar a aritmética dos números naturais é incompleto e, em particular, incapaz de demonstrar a própria consistência.

Ou seja, a razão pela qual a Lógica ocupou-se, e talvez ainda ocupe-se, eminentemente de argumentos matemáticos ou afins é muito mais a persistência de uma tradição do que propriamente lógica; não há nenhuma razão *prima facie* para que a Lógica construa sistemas apenas para inferências de natureza matemática ou semelhante à dessa ciência. Com efeito, boa parte das pesquisas do século XX dedicou-se à investigação da viabilidade de formalizarem-se outros tipos de inferência. Um exemplo são as lógicas modais e a semântica de mundos possíveis de Kripke. Esse paradigma permite falar-se formalmente sobre possibilidade e necessidade, no caso das lógicas aléticas, ou sobre permissibilidade e obrigatoriedade, no caso das deônticas. A compatibilidade entre o sistema D e o discurso do Direito ilustra bem as possibilidades ricas da Lógica para além da Matemática[10].

[10] O sistema modal mais simples é o K, que estende $\mathcal{L}_{\text{LPC}}$ com o operador $\Box$. Isso é feito por meio da adição do axioma K e da regra de inferência denominada "necessitação". O axioma K é $\Box(\alpha \to \beta) \to (\Box\alpha \to \Box\beta)$, em que $\Box$ é um conectivo unário geralmente lido como "é necessário que (seja o caso)". O axioma K estipula a distributividade do novo conectivo sobre a implicação, portanto. A regra de necessitação é $\models_{\text{K}} \alpha \implies \models_{\text{K}} \Box\alpha$. O sistema modal D difere de K, por sua vez, apenas por adicionar-lhe o axioma D, que é $\Box\alpha \to \Diamond\alpha$. Nesse sistema, $\Box$ é reinterpretado como obrigatoriedade (de se fazer algo), e $\Diamond$ é um conectivo definido a partir de $\Box$ interpretado

Uma vez expandidos os horizontes da Lógica, fica imediatamente claro que em campos que não as ciências formais o raciocínio se dá sob incerteza, nas suas mais diversas formas. Tome-se como exemplo um paciente que consulta vários médicos a respeito de seu problema de saúde. Um médico pode dizer-lhe “tome paracetamol à noite”, sentença que contém incerteza porque carece de informações: que quantidade de paracetamol? Outro pode receitar “tome paracetamol e dipirona 25mg à noite”, o que é ambíguo, pois não fica claro a qual medicamento corresponde a posologia. Também aí há incerteza. Pode acontecer ainda que o paciente tenha razões para suspeitar das credenciais do médico, ou ter razões para crer que ele não está dedicando-se plenamente a avaliá-lo, o que impõrá incerteza sobre o que ele receitar.

Isso leva naturalmente à ideia de que uma lógica que pretenda abarcar o discurso empírico (científico ou cotidiano) deve ou mitigar a incerteza, ou abandonar certos preceitos clássicos a fim de poder inferir conclusões razoáveis mesmo na presença de inconsistência:

> *Formal logical systems are supposed to be relevant to the assessment of informal arguments; but the classical logical systems, in which every wff [well formed formula] is either true or else false, seem inappropriate for the assessment of informal arguments with premises and/or conclusions which, because of their vagueness, one hesitates to call either definitely true or definitely false. Once the problem has been put in this way, there seem to be two natural approaches to its solution: to tidy up the vague informal arguments before submitting them to assessment by the standards of classical, 2-valued logic, or, to devise some alternative formal logical system which apply to them more directly (HAACK, 1978, p. 163).*

A primeira estratégia, focada em desfazer as inconsistências, foi implementada principalmente por sistemas de *revisão de crença*, em que proposições inconsistentes são retratadas até que a teoria em questão possa ser tratada classicamente: “the goal of the first approach is to make an inconsistent theory consistent, either by revising it or by representing it by a consistent semantics. So, the main concern there is to avoid contradictions” (DE AMO; CARNIELLI; MARCOS, 2002, p. 68). Na abordagem que renuncia à lógica clássica, destacar-se-ão as lógicas paraconsistentes: “the paraconsistent approach allows reasoning in the presence of inconsistency, and contradictory information can be derived or introduced without trivialization” (DE AMO; CARNIELLI; MARCOS, 2002, p. 68). Este trabalho restringe-se, dados seu escopo e propósitos, à alternativa das lógicas paraconsistentes.

Como visto, as lógicas clássicas distinguem, em suas valorações, proposições verdadeiras de falsas. Um modo de superar a bivalência para incorporar incerteza ao sistema lógico é substituir o conjunto clássico de valores de verdade, $\{\mathtt{t}, \mathtt{f}\}$, por um conjunto com mais de dois valores de verdade, geralmente numéricos. O intuito é que a lógica resultante formalize os graus de crença (de um agente racional ideal) nas proposições. O conjunto de valores de verdade substitutos pode ser finito ou infinito.

como permissibilidade (de se fazer algo). O axioma D afirma que, se uma ação é obrigatória, então é permitido que seja executada (o que realmente soa como um pré-requisito para qualquer legislação razoável).

No caso finito, o conjunto de valores de verdade geralmente é da forma

$$\left\{\frac{n-i}{n} : n \geq 1 \land 0 \leq i \leq n\right\}$$

Tipicamente, 0 e 1 representam a falsidade e a verdade, respectivamente, e os números intermediários representam outros graus de crença. O conjunto de valores de verdade sempre contêm exatamente duas partições: um subconjunto de valores de verdade ditos "distinguidos", e um de "não distinguidos". Um exemplo seria uma lógica com $\{0, {}^1\!/\!_2, 1\}$ como conjunto de valores de verdade, sendo 0 o único valor não distinguido. Pode-se imaginar o valor de verdade intermediário como ignorância, abstenção de juízo, ou ainda incerteza absoluta.

No caso infinito, o conjunto dos números reais, ou um de seus intervalos, funciona como novo contradomínio das valorações, e nesses casos temos, entre outras, as lógicas probabilísticas. Nem toda lógica cujos valores de verdade são números reais é probabilística: é preciso ainda que o valor de verdade das fórmulas moleculares seja determinado conforme o cálculo de probabilidade (a axiomática de Kolmogorov, por exemplo). As lógicas *fuzzy* são exemplos de lógicas com valores de verdade reais que não são probabilísticas.

Uma vez adotados números reais como valores de verdade, o número atribuído a uma proposição é não mais a sua verdade, mas a sua "assertabilidade", isto é, o quão inclinado a asseverá-la um agente racional ideal sentir-se-ia. Ou ainda, o quanto acreditaria nela, ou o quanto estaria disposto a apostar em sua veracidade.

Esse capítulo introduziu o aparato formal e conceitual necessário para a compreensão das demonstrações centrais nesta dissertação, e também expôs os principais resultados de LPC relevantes para esse entendimento. Assim, passa-se agora à análise detalhada das demonstrações dos resultados de trivialização, com o intuito de posteriormente tentar reproduzi-los para outras lógicas, em particular a lógica paraconsistente LFI1.

3 Certos fatos inesperados

> Todas as coisas são canseiras tais, que ninguém as pode exprimir; os olhos não se fartam de ver, nem se enchem os ouvidos de ouvir.
>
> Eclesiastes 1:8 (Almeida Revista e Atualizada).

Uma lógica probabilística pode ser axiomatizada a partir de conceitos semânticos ou de modo puramente sintático. O primeiro método antecede o segundo e é, poder-se-ia argumentar, mais intuitivo. A diferença entre as duas alternativas fica mais evidente ao se examinar uma axiomática semântica. Considerem-se esses axiomas, adaptados de Roeper e Leblanc (1999, p. 6):

AK1 $0 \leq p(\alpha)$;

AK2 $p(\sim(\alpha \wedge \sim\alpha)) = 1$;

AK3 Se $\alpha \wedge \beta$ é uma contradição, então $p(\alpha \vee \beta) = p(\alpha) + p(\beta)$;

AK4 Se α e β são logicamente equivalentes, então $p(\alpha) = p(\beta)$.

Essa axiomática é semântica porque seus dois últimos axiomas são condicionais cujos antecedentes são fatos semânticos, isto é, afirmações a respeito dos valores de verdade das proposições. É também nesse sentido que essas axiomáticas são ditas "não-autônomas"; não são autônomas em relação à semântica da lógica subjacente.

Popper argumentou, em 1938, pela importância de axiomatizações "autônomas", ou seja, independentes de noções semânticas. Isso é particularmente caro a tentativas de mostrar que a teoria da probabilidade é uma generalização da Lógica: se as funções de probabilidade são mais gerais que as valorações e as englobam, deve ser possível construí-las sem recorrer a estas últimas. Abaixo, uma axiomática autônoma, adaptada de Popper por Roeper e Leblanc (1999, p. 6–7):

AP1 $0 \leq p(\alpha)$ (não-negatividade)

AP2 $p(\sim(\alpha \wedge \sim\alpha)) = 1$ (normalidade)

AP3 $p(\alpha) = p(\alpha \wedge \beta) + p(\alpha \wedge \sim\beta)$ (adição)

AP4 $p(\alpha \wedge \beta) \leq p(\beta \wedge \alpha)$ (comutatividade)

AP5 $p(\alpha \wedge (\beta \wedge \gamma)) \leq p((\alpha \wedge \beta) \wedge \gamma)$ (associatividade)

AP6 $p(\alpha) \leq p(\alpha \wedge \alpha)$ (idempotência)

Os teoremas que ficaram conhecidos como “leis do excesso”, bem como os resultados de trivialização de Lewis, serão demonstrados a partir de uma axiomática adaptada do sistema de Lewis (1976), que não é autônoma. Essa axiomática será exposta mais adiante, na Definição 3.0.1; por ora, comentar-se-á em que medida ela difere do sistema em que se baseia. Como já comentado, as funções de probabilidade podem ser definidas em termos de conjuntos ou de proposições lógicas, e as definições são formalmente equivalentes, dado o isomorfismo entre álgebras de conjuntos (Definição 1.2.1) e álgebras booleanas. Assim, os axiomas de Lewis (1976) são uma adaptação do sistema conjuntístico de Kolmogorov (Definição 1.2.2) à LPC. Mais especificamente, eles adaptam o caso finito do sistema de Kolmogorov, isto é, o sistema {K1, K2, K3}. A versão de Lewis (1976) emprega, no entanto, uma semântica baseada em um modelo de Kripke como os empregados nas lógicas modais: a interpretação dessa linguagem fornece, para cada sentença, seu valor de verdade em cada mundo possível. Nesse sentido, duas sentenças são *equivalentes* se são verdadeiras exatamente nos mesmos mundos, e *incompatíveis* se não existe um mundo em que sejam ambas verdadeiras. Uma sentença α *implica* uma sentença β se β é verdadeira em todos os mundos possíveis em que α é verdadeira. Uma sentença é *necessária* se é verdadeira em todos os mundos possíveis, *impossível* se é falsa em todos os mundos possíveis, e *possível* se não é impossível. A modalidade pode ser substituída por noções semânticas[1] não-modais de maneira razoavelmente simples. A equivalência entre duas proposições α e β pode ser expressa por $\alpha \mathrel{=\!\!|\!\!\models} \beta$. Isto é, elas são “equivalentes” se assumir uma como premissa basta para deduzir a outra. Já a incompatibilidade pode ser entendida como $\varnothing \models \sim(\alpha \wedge \beta)$. Isto é, duas proposições são incompatíveis se, e somente se, sua conjunção é uma partícula *bottom*. Finalmente, uma proposição é necessária se equivale, metalogicamente, a uma tautologia qualquer. Essa reformulação dos axiomas de Lewis (1976) torna-os muito semelhantes aos de Fitelson (2008, p. 111).

Comecemos com uma linguagem sintaticamente igual a LPC, isto é, uma linguagem cuja assinatura contém os mesmos símbolos da assinatura de LPC, e somente eles. Deseja-se que sua semântica não seja clássica, mas probabilística, isto é, que incorpore a axiomática de Kolmogorov (Definição 1.2.2). A função de valoração de LPC será substituída por uma função de probabilidade absoluta (unária, portanto) cujo funcionamento será regido pelos seguintes axiomas:

Definição 3.0.1 (Axiomática adaptada de Lewis (1976, p. 299))**.**

Cdmn $0 \leq p(\alpha) \leq 1$ (contradomínio)

Equi Se $\alpha \mathrel{=\!\!|\!\!\models} \beta$, então $p(\alpha) = p(\beta)$ (equivalência)

Disj Se $\varnothing \models \sim(\alpha \wedge \beta)$, então $p(\alpha \vee \beta) = p(\alpha) + p(\beta)$ (disjunção)

Need Se $\varnothing \models \alpha$, então $p(\alpha) = 1$ (necessidade)

[1] Ou sintáticas, já que LPC é uma lógica correta e completa. Seria possível definir equivalência metalinguística entre α e β como $\alpha \dashv\vdash \beta$, por exemplo.

A probabilidade absoluta definirá a probabilidade condicional, uma função binária:

Definição 3.0.2 (Probabilidade condicional)**.**

$$p(\alpha \mid \beta) := \begin{cases} \dfrac{p(\alpha \land \beta)}{p(\beta)} & \text{se } p(\beta) > 0 \\ \text{indefinida} & \text{se } p(\beta) = 0 \end{cases}$$

Seria possível fazer o inverso, isto é, tomar a probabilidade condicional como primitiva e definir a absoluta a partir dela. Essa não foi a opção adotada aqui em parte porque isso tornaria as demonstrações desnecessariamente mais complexas e extensas, em parte porque preferiu-se manter a axiomática o mais parecida com a de Lewis (1976) quanto possível.

Uma vez introduzidos os axiomas, é importante falar sobre condicionalização e independência. A condicionalização é uma operação que gera uma função de probabilidade unária a partir de uma função de probabilidade binária; trata-se de uma forma de *currying*[2], portanto. Pode-se pensar que, se $P_2 = \{p\colon p(\alpha, \beta) = p(\alpha \mid \beta)\}$ for o conjunto de funções de probabilidade binárias que têm como valor a probabilidade de seu primeiro argumento dado seu segundo argumento, e se P_1 for o conjunto de funções de probabilidade unárias, então a condicionalização é uma função $P_2 \mapsto P_1$ que, para cada função da forma $p(\alpha \mid \beta)$, tem como valor a função $p_\beta(\alpha) = p(\alpha \mid \beta)$. Diz-que p_β é *gerada sobre condicionalização* a partir de p: “whenever p(B) is positive, there is a probability function p' such that p'(A) always equals p(A | B); we say that p' comes from p by conditionalizing on B” (LEWIS, 1976, p. 299). É esta, portanto, a definição de condicionalização:

Definição 3.0.3 (Condicionalização)**.** $p_\beta(\alpha) := p(\alpha \mid \beta)$, se $p(\beta) > 0$.

Uma classe de funções de probabilidade está *fechada sob condicionalização* se toda função gerada por condicionalização sobre uma função da classe também pertence à classe (LEWIS, 1976, p. 299).

Independência é uma noção muito importante para a teoria da probabilidade. Formalmente, dois eventos, ou sentenças que descrevem eventos, são independentes se a probabilidade do primeiro condicionado ao segundo é igual à sua probabilidade absoluta:

Definição 3.0.4 (Independência e dependência)**.** α e β são *independentes* se $p(\alpha \mid \beta) = p(\alpha)$. São *dependentes* se não são independentes.

Intuitivamente, pode-se pensar que condicionalizar “não faz diferença” caso os argumentos da função sejam independentes. Assim como acontece com o conceito de probabilidade, o significado da independência depende da visão teórica adotada. Para um frequentista, dois

[2] Também conhecida como *Schönfinkelização*, *currying* é uma operação que reduz funções de aridade igual ou maior a dois a uma sequência de funções unárias. Por exemplo, dada uma função $f\colon \mathbb{N}^3 \mapsto \mathbb{N}$, sua versão reduzida por *currying* será $f'\colon \mathbb{N} \mapsto (\mathbb{N} \mapsto (\mathbb{N} \mapsto \mathbb{N}))$ se, e somente se, for o caso que $f(i, j, k) = f'(i)(j)(k)$, para todos $i, j, k \in \mathbb{N}$.

eventos (ou sentenças que descrevem eventos) são independentes se a frequência de um evento ocorrer quando um segundo evento ocorreu é igual à a mesma sua frequência quando esse outro evento não ocorreu. Em uma perspectiva epistêmica, a independência consiste na ausência de motivação para mudar-se a confiança em uma sentença conforme a confiança em outra sentença. E assim por diante. É fácil entender porque a independência é uma noção importante não só para a fundamentação teórica da teoria da probabilidade, mas também para suas aplicações práticas. Em Estatística, por exemplo, uma questão frequente é, dados dois eventos, saber se eles "têm a ver um com o outro", isto é, saber se são ou não independentes. Por exemplo, ao testar uma nova droga para decidir sua aprovação como medicamento, é importante saber se há dependência, e se houver, quão grande é, entre "João é um usuário da droga" e "João está melhor da doença à qual a droga destina-se".

Uma vez discutidos esses aspectos, procede-se à prova de alguns lemas básicos essenciais para os teoremas posteriores.

3.1 Lemas básicos

Lema 3.1.1 (Conjunção). *$p(\alpha \wedge \beta) = p(\alpha \mid \beta)p(\beta)$, se $p(\beta) > 0$.*

Demonstração. Segue-se imediatamente da Definição 3.0.2. ■

Lema 3.1.2 (Independência). *$p(\alpha \mid \beta) = p(\alpha)$ (isto é, α e β são independentes (Definição 3.0.4) se, e somente se, $p(\alpha \wedge \beta) = p(\alpha)p(\beta)$).*

Demonstração.
Direção ($\Rightarrow$):

(3.1)	$p(\alpha \mid \beta) = p(\alpha)$	hipótese
(3.2)	$\dfrac{p(\alpha \wedge \beta)}{p(\beta)} = p(\alpha)$	Definição 3.0.2, (3.1)
(3.3)	$p(\alpha \wedge \beta) = p(\alpha)p(\beta)$	(3.2)

Direção ($\Leftarrow$):

(3.4)	$p(\alpha \wedge \beta) = p(\alpha)p(\beta)$	hipótese
(3.5)	$p(\alpha \mid \beta) = \dfrac{p(\alpha \wedge \beta)}{p(\beta)}$	Definição 3.0.2
(3.6)	$= \dfrac{p(\alpha)p(\beta)}{p(\beta)}$	(3.4), (3.5)
(3.7)	$= p(\alpha)$	(3.6)

■

Lema 3.1.3 (Complemento)**.** $p(\alpha) = 1 - p({\sim}\alpha)$.

Demonstração.

(3.8)	$p(\alpha \vee {\sim}\alpha) = 1$	Teorema 2.1.8, Need
(3.9)	$p(\alpha) + p({\sim}\alpha) = 1$	Teorema 2.1.10, Disj, (3.8)
(3.10)	$p(\alpha) = 1 - p({\sim}\alpha)$	(3.9)

■

Lema 3.1.4 (Expansão)**.** $p(\alpha) = p(\alpha \wedge \beta) + p(\alpha \wedge {\sim}\beta)$.

Demonstração.

(3.11)	$p(\alpha) = p([\alpha \wedge \beta] \vee [\alpha \wedge {\sim}\beta])$	Teorema 2.1.9, Equi
(3.12)	$= p(\alpha \wedge \beta) + p(\alpha \wedge {\sim}\beta)$	Teorema 2.1.11, Disj, (3.11)

■

Comentário. Os Lemas 3.1.3 e 3.1.4 dependem diretamente dos teoremas clássicos de não-contradição (Teoremas 2.1.10 e 2.1.11), de modo que é esperado que não vigorem em LFI1.

Corolário 3.1.5 (Contradição)**.** $p(\bot) = 0$.

Demonstração. Qualquer contradição é equivalente à negação de qualquer tautologia: $\bot \leftrightarrow {\sim}\top$. Devido a essa equivalência e ao axioma Equi, temos que $p(\bot) = p({\sim}\top)$. Segue-se do Lema 3.1.3 que $p(\bot) = 1 - p(\top)$. Mas $p(\top) = 1$, pelo axioma Need; então, $p(\bot) = 1 - 1 = 0$. ■

Lema 3.1.6 (Probabilidade condicional equivalente)**.** *Se* $\beta \dashv\models \gamma$, *então* $p(\alpha \mid \beta) = p(\alpha \mid \gamma)$.

Demonstração.

(3.13)	$\beta \dashv\models \gamma$	hipótese
(3.14)	$p(\alpha \mid \beta) = \dfrac{p(\alpha \wedge \beta)}{p(\beta)}$	Definição 3.0.2
(3.15)	$= \dfrac{p(\alpha \wedge \gamma)}{p(\gamma)}$	Equi, (3.13), (3.14)
(3.16)	$= p(\alpha \mid \gamma)$	Definição 3.0.2, (3.15)

■

Lema 3.1.7 (Probabilidade condicional tautológica)**.** $p(\alpha \mid \top) = p(\alpha)$.

Demonstração.

(3.17)	$p(\alpha \mid \top) = \dfrac{p(\alpha \wedge \top)}{p(\top)}$	Definição 3.0.2
(3.18)	$= \dfrac{p(\alpha)}{p(\top)}$	Teorema 2.1.4, Equi, (3.17)
(3.19)	$= p(\alpha)$	Need, (3.18)

■

3.2 Lemas intermediários

Lema 3.2.1 (Probabilidade condicional redundante)**.** *Se* $p(\alpha \wedge \beta) > 0$*, então* $p(\alpha \mid \alpha \wedge \beta) = p(\alpha \mid \beta \wedge \alpha) = 1$.

Demonstração.

(3.20)	$p(\alpha \wedge \beta) > 0$	hipótese
(3.21)	$p(\alpha \mid \alpha \wedge \beta) = \dfrac{p(\alpha \wedge (\alpha \wedge \beta))}{p(\alpha \wedge \beta)}$	Definição 3.0.2, (3.20)
(3.22)	$= \dfrac{p(\alpha \wedge \beta)}{p(\alpha \wedge \beta)}$	Teorema 2.1.2, Teorema 2.1.1, Equi, (3.21)
(3.23)	$= 1$	(3.22)
(3.24)	$p(\alpha \mid \beta \wedge \alpha) = 1$	Teorema 2.1.1, Lema 3.1.6, (3.23)

■

Lema 3.2.2 (Disjunção)**.** $p(\alpha \vee \beta) = p(\alpha) + p(\beta) - p(\alpha \wedge \beta)$.

Demonstração.

(3.25)	$p(\beta) = p(\beta \wedge \alpha) + p(\beta \wedge {\sim}\alpha)$	Lema 3.1.4
(3.26)	$p({\sim}\alpha \wedge \beta) = p(\beta) - p(\alpha \wedge \beta)$	Teorema 2.1.1, Equi, (3.25)
(3.27)	$p({\sim}\alpha) = p({\sim}\alpha \wedge \beta) + p({\sim}\alpha \wedge {\sim}\beta)$	Lema 3.1.4
(3.28)	$p({\sim}\alpha \wedge {\sim}\beta) = p({\sim}\alpha) - p({\sim}\alpha \wedge \beta)$	(3.27)
(3.29)	$= p({\sim}\alpha) - [p(\beta) - p(\alpha \wedge \beta)]$	(3.26), (3.28)
(3.30)	$= p({\sim}\alpha) - p(\beta) + p(\alpha \wedge \beta)$	(3.29)
(3.31)	$p(\alpha \vee \beta) = p({\sim}[{\sim}\alpha \wedge {\sim}\beta])$	Teorema 2.1.6, Equi
(3.32)	$= 1 - p({\sim}\alpha \wedge {\sim}\beta)$	Lema 3.1.3, (3.31)
(3.33)	$= 1 - [p({\sim}\alpha) - p(\beta) + p(\alpha \wedge \beta)]$	(3.30), (3.32)
(3.34)	$= 1 - p({\sim}\alpha) + p(\beta) - p(\alpha \wedge \beta)$	(3.33)
(3.35)	$= p(\alpha) + p(\beta) - p(\alpha \wedge \beta)$	Lema 3.1.3, (3.34)

■

Comentário. O passo (3.31) envolve a lei de De Morgan da conjunção (Teorema 2.1.6), que expressa esse conectivo em termos apenas da negação e da disjunção.

Lema 3.2.3 (Implicação material)**.** $p(\alpha \to \beta) = p(\sim\alpha) + p(\beta) - p(\sim\alpha \wedge \beta)$.

Demonstração.

(3.36)	$p(\alpha \to \beta) = p(\sim\alpha \vee \beta)$	Teorema 2.1.5, Equi
(3.37)	$= p(\sim\alpha) + p(\beta) - p(\sim\alpha \wedge \beta)$	Lema 3.2.2, (3.36)

■

Lema 3.2.4 (Equivalência material)**.** *Se* $p(\alpha \leftrightarrow \beta) = 1$, *então* $p(\alpha) = p(\beta)$.

Demonstração.

(3.38)	$p(\alpha \leftrightarrow \beta) = 1$	hipótese
(3.39)	$p([\alpha \to \beta] \wedge [\beta \to \alpha]) = 1$	Definição de $\leftrightarrow$, Equi, (3.38)
(3.40)	$p(\alpha \to \beta) = 1$	(3.39)
(3.41)	$p(\beta \to \alpha) = 1$	(3.39)
(3.42)	$p(\sim\alpha) + p(\beta) - p(\sim\alpha \wedge \beta) = 1$	Lema 3.2.3, (3.40)
(3.43)	$p(\beta) = 1 - p(\sim\alpha) + p(\sim\alpha \wedge \beta)$	(3.42)
(3.44)	$p(\beta) = p(\beta \wedge \alpha) + p(\beta \wedge \sim\alpha)$	Lema 3.1.4
(3.45)	$p(\beta \wedge \alpha) = 1 - p(\sim\alpha)$	Teorema 2.1.1, (3.43), (3.44)
(3.46)	$p(\beta \wedge \alpha) = p(\alpha)$	Lema 3.1.3, (3.45)
(3.47)	$p(\sim\beta) + p(\alpha) - p(\sim\beta \wedge \alpha) = 1$	Lema 3.2.3, (3.41)
(3.48)	$p(\alpha) = 1 - p(\sim\beta) + p(\sim\beta \wedge \alpha)$	(3.47)
(3.49)	$p(\alpha) = p(\alpha \wedge \beta) + p(\alpha \wedge \sim\beta)$	Lema 3.1.4
(3.50)	$p(\alpha \wedge \beta) = 1 - p(\sim\beta)$	Teorema 2.1.1, (3.48), (3.49)
(3.51)	$p(\alpha \wedge \beta) = p(\beta)$	Lema 3.1.3, (3.50)
(3.52)	$p(\alpha) = p(\beta)$	Teorema 2.1.1, (3.46), (3.51)

■

Lema 3.2.5. *Se* $p(\alpha) = 1$, *então* $p(\sim\alpha \wedge \beta) = p(\beta \wedge \sim\alpha) = 0$.

Demonstração.

(3.53)	$p(\alpha) = 1$	hipótese
(3.54)	$p(\sim\alpha) = 0$	Lema 3.1.3, (3.53)
(3.55)	$p(\sim\alpha) = p(\sim\alpha \wedge \beta) + p(\sim\alpha \wedge \sim\beta)$	Lema 3.1.4
(3.56)	$0 = p(\sim\alpha \wedge \beta) + p(\sim\alpha \wedge \sim\beta)$	(3.54), (3.55)
(3.57)	$p(\sim\alpha \wedge \beta) = p(\sim\alpha \wedge \sim\beta) = 0$	Cdmn, (3.56)
(3.58)	$p(\beta \wedge \sim\alpha) = p(\sim\beta \wedge \sim\alpha) = 0$	Teorema 2.1.1, Equi, (3.57)

■

Corolário 3.2.6. *Se* $p(\alpha) = 0$*, então* $p(\alpha \wedge \beta) = p(\beta \wedge \alpha) = 0$*.*

Demonstração. Análoga à demonstração do Lema 3.2.5. ■

Comentário. A contrapositiva do Corolário 3.2.6 será especialmente útil: se $p(\alpha \wedge \beta) > 0$, então $p(\alpha) > 0$[3].

Corolário 3.2.7. *Se* α *é uma contradição, então* $p(\alpha) = p(\alpha \wedge \beta) = p(\beta \wedge \alpha) = 0$*.*

Demonstração. Segue-se imediatamente dos Corolários 3.1.5 e 3.2.6 (e, é claro, da comutatividade da conjunção (Teorema 2.1.1) e do axioma Equi). ■

Lema 3.2.8. *Se* $p(\alpha) = 1$*, então* $p(\alpha \wedge \beta) = p(\beta) = p(\beta \wedge \alpha)$*.*

Demonstração.

(3.59)	$p(\alpha) = 1$	hipótese
(3.60)	$p(\sim\alpha) = 0$	Lema 3.1.3, (3.59)
(3.61)	$p(\alpha \wedge \beta) = 1 - p(\alpha \wedge \sim\beta)$	Lema 3.1.4, (3.59)
(3.62)	$= p(\sim[\alpha \wedge \sim\beta])$	Lema 3.1.3, (3.61)
(3.63)	$= p(\sim\alpha \vee \beta)$	Teorema 2.1.6, Teorema 2.1.12, Equi, (3.62)
(3.64)	$= p(\sim\alpha) + p(\beta) - p(\sim\alpha \wedge \beta)$	Lema 3.2.2, (3.63)
(3.65)	$= p(\beta) - p(\sim\alpha \wedge \beta)$	(3.60), (3.64)
(3.66)	$= p(\beta)$	Lema 3.2.5, (3.59), (3.65)
(3.67)	$p(\beta \wedge \beta) = p(\beta)$	Teorema 2.1.1, Equi, (3.66)

■

3 Dado o axioma Cdmn, $p(\alpha) \neq 0$ equivale a $p(\alpha) > 0$.

Lema 3.2.9 (Conjunção tautológica)**.** *Se* $p(\alpha \wedge \beta) = 1$*, então* $p(\alpha) = p(\beta) = 1$.

Demonstração.

(3.68)	$p(\alpha \wedge \beta) = 1$	hipótese
(3.69)	$p(\alpha) = p(\alpha \wedge \beta) + p(\alpha \wedge {\sim}\beta)$	Lema 3.1.4
(3.70)	$= 1 + 0$	Cdmn, (3.68), (3.69)
(3.71)	$p(\beta) = p(\beta \wedge \alpha) + p(\beta \wedge {\sim}\alpha)$	Lema 3.1.4
(3.72)	$= 1 + 0$	Cdmn, Teorema 2.1.1, Equi, (3.68), (3.71)

■

Lema 3.2.10 (Probabilidade condicional contraditória)**.** $p(\alpha \mid {\sim}\alpha \wedge \beta) = p(\alpha \mid \beta \wedge {\sim}\alpha) = 0$.

Demonstração.

(3.73)	$p(\alpha \mid {\sim}\alpha \wedge \beta) = \dfrac{p(\alpha \wedge [{\sim}\alpha \wedge \beta])}{p({\sim}\alpha \wedge \beta)}$	Definição 3.0.2
(3.74)	$= \dfrac{p(\perp \wedge \beta)}{p(\alpha \wedge \beta)}$	Teorema 2.1.2, Teorema 2.1.10, Equi, (3.73)
(3.75)	$= 0$	Corolário 3.2.7, (3.74)
(3.76)	$p(\alpha \mid \beta \wedge {\sim}\alpha) = 0$	Teorema 2.1.1, Lema 3.1.6, (3.75)

■

Demonstrados esses lemas, passa-se a um teorema importante para o entendimento das leis do excesso e dos resultados de trivialização subsequentes.

Teorema 3.2.11 (Teorema da probabilidade total)**.** $p(\alpha) = p(\alpha \mid \beta)p(\beta) + p(\alpha \mid {\sim}\beta)p({\sim}\beta)$*, se* $0 < p(\beta) < 1$.

Demonstração.

(3.77)	$0 < p(\beta) < 1$	hipótese
(3.78)	$0 < p({\sim}\beta) < 1$	Lema 3.1.3, (3.77)
(3.79)	$p(\alpha) = p(\alpha \wedge \beta) + p(\alpha \wedge {\sim}\beta)$	Lema 3.1.4
(3.80)	$= p(\alpha \mid \beta)p(\beta) + p(\alpha \mid {\sim}\beta)p({\sim}\beta)$	Lema 3.1.1, (3.79), (3.78), (3.77)

■

3.3 Leis do excesso

> Popper's insight of 1938 that the unconditional probability of a condicional statement of the form 'If A, then B' is normally different from the condicional probability of B given A, has become a commonplace among those philosophers who use probability theory for their research.
>
> Dorn (1992–93, p. 3).

No final dos anos 1930, Popper estava axiomatizando a teoria da probabilidade de modo distinto daquele de Kolmogorov e, até onde se sabe, ignorava a existência do trabalho deste último. Em 1938, ele já havia, segundo seu próprio relato, demonstrado alguns resultados envolvendo a relação entre $p(\alpha \to \beta)$ e $p(\beta \mid \alpha)$ (DORN, 1992–93, p. 4). As primeiras publicações sobre esse assunto viriam apenas 30 anos depois, no entanto. Em 1963, Popper publicou seu "Conjecturas e refutações" e, embora tanto sua escrita quanto sua notação dificultem a compreensão das passagens relevantes, ele contém a seguinte equação, cuja notação foi adaptada às nossas convenções: $p(\beta \to \alpha) - p(\alpha \mid \beta) = [1 - p(\alpha \mid \beta)][1 - p(b)]$ (DORN, 1992–93, p. 7). A segunda aparição das chamadas "leis do excesso" viriam em 1966, em um apêndice de "Logik der Forschung" (*Logic of Scientific Discovery*) que não foi reproduzido nas traduções para o inglês da obra (DORN, 1992–93, p. 7–10). É plausível supor que essas primeiras publicações de Popper não repercutiram rapidamente na literatura, em parte devido ao modo obscuro com que foram registradas pelo autor. Seja como for, parece seguro afirmar que as "leis do excesso" atribuíveis a ele são as que se seguem[4].

3.3.1 Primeira lei do excesso

Teorema 3.3.1 (Primeira lei do excesso)**.** *$p(\alpha \to \beta) \geq p(\beta \mid \alpha)$, se $p(\alpha) > 0$.*

[4] Dados seus propósitos, esta dissertação não almeja tratar em detalhes da história das leis do excesso e de sua consolidação na pesquisa sobre a intergace entre Teoria da probabilidade e Lógica. Indica-se Dorn (1992–93) como ponto de partida para os interessados em aprofundarem-se nesse tema.

Demonstração.

(3.81)	$p(\alpha) > 0$	hipótese
(3.82)	$p(\alpha \to \beta) = p({\sim}\alpha \vee \beta)$	Teorema 2.1.5, Equi
(3.83)	$= p({\sim}\alpha \vee (\beta \wedge \alpha))$	Teorema 2.1.8, Equi, (3.82)
(3.84)	${\sim}\alpha \wedge (\beta \wedge \alpha) \leftrightarrow \perp$	Teorema 2.1.2, Teorema 2.1.10
(3.85)	$= p({\sim}\alpha) + p(\beta \wedge \alpha)$	Disj, (3.84), (3.83)
(3.86)	$= p({\sim}\alpha) + p(\beta \mid \alpha)p(\alpha)$	Lema 3.1.1, (3.81), (3.85)
(3.87)	$\geq p({\sim}\alpha)p(\beta \mid \alpha) + p(\beta \mid \alpha)p(\alpha)$	(3.86)
(3.88)	$\geq p(\beta \mid \alpha)[p({\sim}\alpha) + p(\alpha)]$	(3.87)
(3.89)	$\geq p(\beta \mid \alpha)$	Lema 3.1.3, (3.88)

■

3.3.2 Segunda lei do excesso

Teorema 3.3.2 (Segunda lei do excesso)**.** *$p(\alpha \to \beta) = p(\beta \mid \alpha)$ se, e somente se, $p({\sim}\alpha) = 0$ ou $p(\beta \mid \alpha) = 1$.*

Demonstração.

Direção ($\Leftarrow$), caso $p({\sim}\alpha) = 0$:

(3.90)	$p({\sim}\alpha) = 0$	hipótese
(3.91)	$p(\alpha) = 1$	Lema 3.1.3, (3.90)
(3.92)	$p(\alpha \to \beta) = p({\sim}\alpha \vee \beta)$	Teorema 2.1.5, Equi
(3.93)	$= p({\sim}\alpha \vee (\beta \wedge \alpha))$	Teorema 2.1.8, Equi, (3.92)
(3.94)	$= p({\sim}\alpha) + p(\beta \wedge \alpha)$	Teorema 2.1.1, Teorema 2.1.2, Teorema 2.1.10, Disj, (3.93)
(3.95)	$= p({\sim}\alpha) + p(\beta \mid \alpha)p(\alpha)$	Lema 3.1.1, (3.91), (3.94)
(3.96)	$= 0 + p(\beta \mid \alpha)1$	(3.90), (3.91), (3.95)
(3.97)	$= p(\beta \mid \alpha)$	(3.96)

Direção ($\Leftarrow$), caso $p(\beta \mid \alpha) = 1$:

(3.98)	$p(\beta \mid \alpha) = 1$	hipótese
(3.99)	$p(\alpha \to \beta) = p(\sim\alpha \lor (\beta \land \alpha))$	Teorema 2.1.5, Equi
(3.100)	$= p(\sim\alpha) + p(\beta \land \alpha)$	Teorema 2.1.1, Teorema 2.1.2, Disj, (3.99)
(3.101)	$= p(\sim\alpha) + p(\beta \mid \alpha)p(\alpha)$	Lema 3.1.1, (3.98), (3.100)
(3.102)	$= p(\sim\alpha) + p(\alpha)$	(3.98), (3.101)
(3.103)	$= 1$	Lema 3.1.3, (3.102)
(3.104)	$= p(\beta \mid \alpha)$	(3.98), (3.103)

Direção ($\Rightarrow$):

(3.105)	$p(\alpha \to \beta) = p(\beta \mid \alpha)$	hipótese
(3.106)	$p(\alpha) > 0$	(3.105)
(3.107)	$p(\sim\alpha) < 1$	Lema 3.1.3, (3.106)
(3.108)	$p(\alpha \to \beta) = p(\sim\alpha \lor (\alpha \land \beta))$	Teorema 2.1.5, Teorema 2.1.8, Equi, (3.105)
(3.109)	$p(\sim\alpha \lor (\alpha \land \beta)) = p(\beta \mid \alpha)$	(3.105), (3.108)
(3.110)	$p(\sim\alpha) + p(\alpha \land \beta) = p(\beta \mid \alpha)$	Equi, Teorema 2.1.2, Disj, (3.109)
(3.111)	$= \dfrac{p(\beta \land \alpha)}{p(\alpha)}$	Definição 3.0.2, (3.110)

I. Caso $p(\sim\alpha) = 0$:

(3.112)	$p(\sim\alpha) = 0$	hipótese
(3.113)	$p(\alpha) = 1$	Lema 3.1.3, (3.112)
(3.114)	$0 + p(\alpha \land \beta) = \dfrac{p(\beta \land \alpha)}{1}$	(3.111), (3.112), (3.113)

O caso $p(\sim\alpha) = 0$ é possível, portanto (Teorema 2.1.1, Equi).

II. Caso $p(\sim\alpha) > 0$:

(3.115)	$0 < p(\sim\alpha) < 1$	hipótese, (3.107)
(3.116)	$0 < p(\alpha) < 1$	Lema 3.1.3, (3.115)

i. Caso $p(\alpha \land \beta) = 0$:

(3.117)	$p(\alpha \land \beta) = 0$	hipótese (absurda)
(3.118)	$p(\sim\alpha) + 0 = \dfrac{0}{p(\alpha)}$	(3.111), (3.117)
(3.119)	$p(\sim\alpha) = 0$	(3.118)

O caso II.i é absurdo, pois (3.119) contradiz (3.115).

ii. Caso $p(\alpha \wedge \beta) = 1$:

(3.120) $\quad p(\alpha \wedge \beta) = 1$ $\quad$ hipótese (absurda)

(3.121) $\quad p(\sim\alpha) + 1 = \dfrac{1}{p(\alpha)}$ $\quad$ (3.111), (3.120)

Seja $x = p(\sim\alpha)$:

(3.122) $\quad x + 1 = \dfrac{1}{1 - x}$ $\quad$ (3.121)

(3.123) $\quad x - x^2 + 1 - x = 1$ $\quad$ (3.122)

(3.124) $\quad x = 0$ $\quad$ (3.123)

O caso II.ii é absurdo, pois (3.124) contradiz (3.115).

iii. Caso $0 < p(\alpha \wedge \beta) < 1$:

Sejam $x = p(\sim\alpha)$ e $y = p(\alpha \wedge \beta)$:

(3.125) $\quad x + y = \dfrac{y}{1 - x}$ $\quad$ (3.111)

(3.126) $\quad x^2 + xy - x = 0$ $\quad$ (3.125)

(3.127) $\quad x(x + y - 1) = 0$ $\quad$ (3.126)

Assim, $x = 0$ e/ou $x + y - 1 = 0$. $x = 0$ é absurdo, pois contradiria (3.115). Logo:

(3.128) $\quad p(\sim\alpha) + p(\alpha \wedge \beta) - 1 = 0$ $\quad$ (3.115), (3.127)

(3.129) $\quad \dfrac{p(\beta \wedge \alpha)}{p(\alpha)} = 1$ $\quad$ (3.111), (3.128)

(3.130) $\quad p(\beta \mid \alpha) = 1$ $\quad$ Definição 3.0.2, (3.129)

Isto é, se $p(\alpha \rightarrow \beta) = p(\beta \mid \alpha)$, então ou $p(\sim\alpha) = 0$, ou $p(\beta \mid \alpha) = 1$. ■

Os Teoremas 3.3.1 e 3.3.2 evidenciam que a probabilidade da implicação material $\rightarrow$ não pode ser igualada à respectiva probabilidade condicional. Mas persiste a intuição de que as probabilidades condicionais expressam alguma relação lógica formalizável por um conectivo; um conectivo *probabilístico*, portanto, na terminologia de Lewis (1976). Tal conectivo será representado, aqui, pelo símbolo $\rightsquigarrow$:

Definição 3.3.1 (Condicional probabilístico)**.** $p(\alpha \rightsquigarrow \beta) := p(\beta \mid \alpha)$, se $p(\alpha) > 0$.

Se a Definição 3.3.1 vigora para uma determinada função de probabilidade p, então $\rightsquigarrow$ é um *condicional probabilístico para p*. Se C_p é uma classe de funções de probabilidades e $\rightsquigarrow$ é um condicional probabilístico para cada um de seus elementos, então $\rightsquigarrow$ é um *condicional probabilístico para* C_p. $\rightsquigarrow$ é um *condicional probabilístico universal* se é um condicional para toda função probabilística, isto é, se a Definição 3.3.1 vale para quaisquer p, α e β (LEWIS, 1976, p. 299). Será útil convencionar que proposições da forma $\alpha \rightsquigarrow \beta$ serão chamadas de *implicações condicionais* (em contraste com as implicações materiais, que são proposições da forma $\alpha \rightarrow \beta$).

O fato de que $\rightsquigarrow$ não pode simplesmente ser a implicação material $\rightarrow$, dado o Teorema 3.3.2, é tomado em Lewis (1976) como óbvio. O autor reconhece-o em um comentário bastante breve e não o credita nem a Popper nem a outro autor qualquer: "[...] of course the ordinary indicative conditional $\alpha \rightsquigarrow \beta$ cannot be the truth-functional conditional $\alpha \to \beta$. $p(\alpha \to \beta)$ and $p(\beta \mid \alpha)$ are equal only in certain extreme cases" (LEWIS, 1976, p. 298)[5].

A hipótese de existir o condicional $\rightsquigarrow$ da Definição 3.3.1 revelar-se-á falsa, como será visto adiante. Para verificar que assim é, a hipótese será mantida, a fim de que resulte em um absurdo. Por ora, cabe comentar que a inviabilidade da Definição 3.3.1 não é óbvia. Se ela parecer sê-lo, quando os chamados "resultados de trivialização" forem detalhadamente examinados adiante, é apenas porque qualquer fato parece óbvio depois de constatado por outra pessoa. Lewis (1976) inclui, entre os que aceitaram a existência de um conectivo como $\rightsquigarrow$: Richard Jeffrey, Brian Ellis, e Robert Stalnaker.

Uma vez definido o conectivo $\rightsquigarrow$, que supostamente formaliza logicamente a probabilidade condicional, provam-se um lema e um corolário necessários para a demonstração da terceira lei do excesso.

Lema 3.3.3 (Probabilidade condicional composta (LEWIS, 1976, p. 299))**.** *Se* $p(\alpha \wedge \gamma) > 0$, *então* $p(\alpha \rightsquigarrow \beta \mid \gamma) = p(\beta \mid \alpha \wedge \gamma)$.

Demonstração.

(3.131) $p(\alpha \wedge \gamma) > 0$ — hipótese

(3.132) $p(\alpha) > 0$ — Contrapositiva do Corolário 3.2.6, (3.131)

(3.133) $p(\gamma) > 0$ — Contrapositiva do Corolário 3.2.6, (3.131)

(3.134)

$p(\alpha \rightsquigarrow \beta \mid \gamma) = p_\gamma(\alpha \rightsquigarrow \beta)$ — Definição 3.0.3, (3.133)

(3.135) $= p_\gamma(\beta \mid \alpha)$ — Definição 3.3.1, (3.134)

(3.136) $= \dfrac{p_\gamma(\beta \wedge \alpha)}{p_\gamma(\alpha)}$ — Definição 3.0.2, (3.135)

(3.137) $= \dfrac{p(\beta \wedge \alpha \mid \gamma)}{p(\alpha \mid \gamma)}$ — Definição 3.0.3, (3.136)

(3.138) $= \dfrac{p((\beta \wedge \alpha) \wedge \gamma)}{p(\gamma)} \div \dfrac{p(\alpha \wedge \gamma)}{p(\gamma)}$ — Definição 3.0.2, (3.137)

(3.139) $= \dfrac{p(\beta \wedge (\alpha \wedge \gamma))}{p(\alpha \wedge \gamma)}$ — Teorema 2.1.2, Equi, (3.131), (3.138)

(3.140) $= p(\beta \mid \alpha \wedge \gamma)$ — Definição 3.0.2, (3.139)

■

[5] A notação de Lewis (1976) foi adaptada, nessa citação e nas demais, às convenções adotadas nesta dissertação; considerou-se mais importante manter a uniformidade da simbologia que reproduzir o texto original *ipsis litteris*. Essa foi a única adaptação feita, no entanto.

Lema 3.3.4 (Condicionalização composta)**.** *Se* $p(\alpha \wedge \beta) > 0$*, então* $p_\beta(\alpha \rightsquigarrow \gamma) = p(\gamma \mid \alpha \wedge \beta) = p(\gamma \mid \beta \wedge \alpha)$.

Demonstração.

(3.141)	$p_\beta(\alpha \rightsquigarrow \gamma) = p_\beta(\gamma \mid \alpha)$	Definição 3.3.1
(3.142)	$= \dfrac{p_\beta(\gamma \wedge \alpha)}{p_\beta(\alpha)}$	Definição 3.0.2, (3.141)
(3.143)	$= p(\gamma \wedge \alpha \mid \beta) \div p(\alpha \mid \beta)$	Definição 3.0.3, (3.142)
(3.144)	$= \dfrac{p((\gamma \wedge \alpha) \wedge \beta)}{p(\beta)} \div \dfrac{p(\alpha \wedge \beta)}{p(\beta)}$	Definição 3.0.2, (3.143)
(3.145)	$= \dfrac{p(\gamma \wedge (\alpha \wedge \beta))}{p(\alpha \wedge \beta)}$	Teorema 2.1.2, Equi, (3.144)
(3.146)	$= p(\gamma \mid \alpha \wedge \beta)$	Definição 3.0.2, (3.144)
(3.147)	$= p(\gamma \mid \beta \wedge \alpha)$	Teorema 2.1.1, Lema 3.1.6, (3.146)

∎

Lema 3.3.5 (Exportação/importação)**.** *Se* $p(\alpha \wedge \beta) > 0$ *e* $p(\alpha \wedge \gamma) > 0$*, então* $p(\alpha \rightsquigarrow [\beta \rightsquigarrow \gamma]) = p((\alpha \wedge \beta) \rightsquigarrow \gamma) = p((\beta \wedge \alpha) \rightsquigarrow \gamma)$.

Demonstração.

(3.148)	$p(\alpha \wedge \beta) > 0$	hipótese
(3.149)	$p(\alpha \wedge \gamma) > 0$	hipótese
(3.150)	$p(\alpha) > 0$	Contrapositiva do Corolário 3.2.6, (3.149)
(3.151)	$p(\alpha \rightsquigarrow [\beta \rightsquigarrow \gamma]) = p(\beta \rightsquigarrow \gamma \mid \alpha)$	Definição 3.3.1, (3.150)
(3.152)	$= p(\gamma \mid \beta \wedge \alpha)$	Lema 3.3.3, (3.148), (3.151)
(3.153)	$= p((\beta \wedge \alpha) \rightsquigarrow \gamma)$	Definição 3.3.1, (3.152)
(3.154)	$= p(\gamma \mid \alpha \wedge \beta)$	Teorema 2.1.1, Lema 3.1.6, (3.152)
(3.155)	$= p((\alpha \wedge \beta) \rightsquigarrow \gamma)$	Definição 3.3.1, (3.154)

∎

Comentário. $[\alpha \rightarrow (\beta \rightarrow \gamma)] \leftrightarrow [(\alpha \wedge \beta) \rightarrow \gamma]$ é um teorema de LPC conhecido por "regra de exportação/importação", de modo que é interessante notar que o condicional $\rightsquigarrow$, se existisse, seria relativamente semelhante à implicação material.

3.3.3 Terceira lei do excesso

Teorema 3.3.6 (Terceira lei do excesso). *$p(\alpha \rightsquigarrow \beta \mid \alpha \rightarrow \beta) = 1$, se $p(\alpha \wedge \beta) > 0$.*

Demonstração.

(3.156)	$p(\alpha \wedge \beta) > 0$	
(3.157)	$p(\alpha \rightsquigarrow \beta \mid \alpha \rightarrow \beta) = p([\alpha \rightarrow \beta] \rightsquigarrow [\alpha \rightsquigarrow \beta])$	Definição 3.3.1
(3.158)	$= p([[\alpha \rightarrow \beta] \wedge \alpha] \rightsquigarrow \beta)$	Lema 3.3.5, (3.157)
(3.159)	$= p(\beta \mid [\alpha \rightarrow \beta] \wedge \alpha)$	Definição 3.3.1, (3.158)

Considerando-se que $[(\alpha \rightarrow \beta) \wedge \alpha] \leftrightarrow (\beta \wedge \alpha)$:

(3.160)	$= p(\beta \mid \beta \wedge \alpha)$	Teorema 2.1.7, Equi, (3.159)
(3.161)	$= 1$	Lema 3.2.1, (3.160)

■

Lema 3.3.7. *Se $p(\alpha \wedge \beta) > 0$, então $p(\alpha \rightsquigarrow \beta) \geq p(\alpha \rightarrow \beta)$.*

Demonstração.

(3.162)	$p(\alpha \wedge \beta) > 0$	hipótese
(3.163)	$p(\alpha \rightsquigarrow \beta) = p(\alpha \rightsquigarrow \beta \wedge \alpha \rightarrow \beta) + p(\alpha \rightsquigarrow \beta \wedge {\sim}(\alpha \rightarrow \beta))$	Lema 3.1.4
(3.164)	$p(\alpha \rightsquigarrow \beta) \geq p(\alpha \rightsquigarrow \beta \wedge \alpha \rightarrow \beta)$	Cdmn, (3.163)
(3.165)	$p(\alpha \rightsquigarrow \beta \wedge \alpha \rightarrow \beta) = p(\alpha \rightsquigarrow \beta \mid \alpha \rightarrow \beta)p(\alpha \rightarrow \beta)$	Lema 3.1.1
(3.166)	$p(\alpha \rightsquigarrow \beta) \geq p(\alpha \rightsquigarrow \beta \mid \alpha \rightarrow \beta)p(\alpha \rightarrow \beta)$	(3.164), (3.165)
(3.167)	$p(\alpha \rightsquigarrow \beta \mid \alpha \rightarrow \beta) = 1$	Teorema 3.3.6, (3.162)
(3.168)	$p(\alpha \rightsquigarrow \beta) \geq p(\alpha \rightarrow \beta)$	(3.166), (3.167)

■

Comentário. Já é possível entrever que esse último lema será combinado à primeira lei do excesso (Teorema 3.3.1) para obter-se a igualdade entre $p(\alpha \rightsquigarrow \beta)$ e $p(\alpha \rightarrow \beta)$.

Demonstrados os teoremas, lemas e corolários anteriores, pode-se passar, enfim, aos resultados de trivialização de Lewis (1976) e de Fitelson (2015).

3.4 Resultados de trivialização

3.4.1 Os primeiros: Lewis (1976)

> Our question, therefore, is whether the indicative conditional might have one fixed interpretation that makes it a probability conditional for the entire class of all those probability functions that represent possible systems of beliefs.
>
> Lewis (1976, p. 302).

Lewis (1976) menciona explicitamente, inclusive por meio da divisão das seções do artigo, dois resultados ditos "de trivialização". A análise desta dissertação, que divide esse texto em três resultados de trivialização, segue a de Milne (2003, p. 300).

Teorema 3.4.1 (Primeiro resultado de trivialização (LEWIS, 1976, p. 300))**.** *Se* $p(\alpha \wedge \beta) > 0$ *e* $p(\alpha \wedge {\sim}\beta) > 0$*, então* $p(\alpha \rightsquigarrow \beta) = p(\beta)$.

Demonstração.

(3.169)	$p(\alpha \wedge \beta) > 0$	hipótese
(3.170)	$p(\alpha \wedge {\sim}\beta) > 0$	hipótese
(3.171)	$p(\beta) > 0$	Contrapositiva do Corolário 3.2.6, (3.169)
(3.172)	$p({\sim}\beta) > 0$	Contrapositiva do Corolário 3.2.6, (3.170)
(3.173)	$p(\alpha \rightsquigarrow \beta \mid \beta) = p(\beta \mid \alpha \wedge \beta)$	Lema 3.3.3, (3.169)
(3.174)	$= 1$	Lema 3.2.1, (3.173)
(3.175)	$p(\alpha \rightsquigarrow \beta \mid {\sim}\beta) = p(\beta \mid \alpha \wedge {\sim}\beta)$	Lema 3.3.3, (3.170)
(3.176)	$= 0$	Lema 3.2.10, (3.175)
(3.177)	$p(\gamma) = p(\gamma \mid \beta)p(\beta) + p(\gamma \mid {\sim}\beta)p({\sim}\beta)$	Teorema 3.2.11, (3.171), (3.172)

Seja $\gamma = \alpha \rightsquigarrow \beta$:

(3.178)	$p(\alpha \rightsquigarrow \beta) = 1p(\beta) + 0p({\sim}\beta)$	(3.174), (3.176), (3.177)
(3.179)	$= p(\beta)$	(3.178)

■

Lema 3.4.2. $p(\alpha \rightsquigarrow \beta) = p(\alpha \rightarrow \beta)$*, se* $p(\alpha \wedge \beta) > 0$.

Demonstração.

(3.180)	$p(\alpha \wedge \beta) > 0$	hipótese
(3.181)	$p(\alpha \wedge \beta) = p(\beta \mid \alpha)p(\alpha)$	Lema 3.1.1
(3.182)	$p(\beta \mid \alpha)p(\alpha) > 0$	(3.180), (3.181)
(3.183)	$p(\beta \mid \alpha) > 0$	Cdmn, (3.182)
(3.184)	$p(\alpha) > 0$	Cdmn, (3.182)
(3.185)	$p(\alpha \rightarrow \beta) \geq p(\beta \mid \alpha)$	Teorema 3.3.1, (3.184)
(3.186)	$p(\alpha \rightsquigarrow \beta) \geq p(\alpha \rightarrow \beta)$	Lema 3.3.7, (3.180)
(3.187)	$p(\alpha \rightarrow \beta) \geq p(\alpha \rightsquigarrow \beta)$	Definição 3.3.1, (3.184), (3.185)
(3.188)	$p(\alpha \rightarrow \beta) = p(\alpha \rightsquigarrow \beta)$	(3.186), (3.187)

■

Assim, assumir a igualdade $p(\alpha \rightsquigarrow \beta) = p(\beta \mid \alpha)$ (Definição 3.3.1) implica que $p(\beta \mid \alpha) = p(\beta)$ (Teorema 3.4.1). Isso significa que duas proposições quaisquer, α e β, são independentes quanto a qualquer função de probabilidade que atenda às definições e axiomas anteriores, o que é absurdo. Considere-se, por exemplo, o lançamento de um dado justo de seis faces. Seja α a proposição "o resultado do lançamento é par", e β a proposição "o resultado do lançamento é a face seis". Então $p(\beta) = 1/6$, mas $p(\alpha \rightsquigarrow \beta) = p(\beta \mid \alpha) = 1/3$, o que mostra que os eventos correspondentes a α e β são dependentes (pela Definição 3.0.4), o que contradiz o Teorema 3.4.1. Há um problema, portanto, ou na formulação desse contraexemplo ou nas premissas do teorema. Como esse contraexemplo é indubitavellmente correto, e como muitos outros similares podem ser elencados, é forçoso admitir que o problema deve-se às fundações do Teorema 3.4.1.

Cabe ainda notar que a impossibilidade de haver um condicional probabilístico universal não impossibilita a existência de um condicional probabilístico para a classe que contém exatamente todas as funções de probabilidade que representam crenças de agentes racionais. Se confirmada, essa possibilidade restauraria, embora somente para essa classe de funções, a igualdade entre a probabilidade do condicional probabilístico $\rightsquigarrow$ e a probabilidade condicional:

> *But all is not yet lost for the thesis that probabilities of conditionals are conditional probabilities. A much less than universal probability conditional might be good enough. Our task, after all, concerns subjective probability: probability functions used to represent people's systems of belief (LEWIS, 1976, p. 301).*

Essa observação é particularmente importante para Lewis (1976) porque a concepção de probabilidade nesse trabalho é de probabilidade subjetiva: a tese de Lewis é que incorporar a teoria da probabilidade à semântica de uma lógica implica substituir a noção de valor de verdade pela de *assertabilidade*. A assertabilidade de uma proposição, para um determinado agente racional em um dado momento, é a segurança que ele sente em afirmá-la, e pode ser quantificada por meio de uma função de probabilidade subjetiva[6]: "he [um falante fidedigno] deems it permissible to assert that α only if $p(\alpha)$ is sufficiently close to 1, where p is the probability function that represents his system of degrees of belief at the time. Assertability goes by subjective probability" (LEWIS, 1976, p. 297).

Mas Lewis (1976) argumenta que deve haver um único condicional probabilístico $\rightsquigarrow$ para a classe de funções de probabilidade subjetiva que representam crenças de agentes racionais. Isto é, não pode ser o caso que cada função de probabilidade subjetiva (ou seja, cada sistema de crenças de uma pessoa em um certo instante) iguale $p(\alpha \rightsquigarrow \beta)$ a $p(\beta \mid \alpha)$ a seu próprio modo. Do contrário, seria impossível explicar como uma pessoa pode mudar de opinião com o passar do tempo, ou como duas pessoas podem entenderem-se a ponto de conseguirem debater sobre a verdade de uma proposição (LEWIS, 1976, p. 301): "our question, therefore, is whether the indicative conditional might have one fixed interpretation that makes it a probability conditional for the entire class of all those probability functions that represent possible systems of beliefs" (LEWIS, 1976, p. 302). Ou seja, dada a interpretação da probabilidade adotada em Lewis (1976), validar a Definição 3.3.1 e todos os resultados que dependem dela exige apenas restringir seu escopo à classe das funções de probabilidades que correspondem a sistemas de crenças de agentes racionais: "we need not assume, and indeed it seems rather implausible, that any probability function whatever represents a system of beliefs that it is possible for someone to have. We might set aside those probability functions that do not" (LEWIS, 1976, p. 301). Como mostra o segundo resultado de trivialização, no entanto, nem mesmo essa restrição pode preservar a igualdade intuitiva entre $p(\alpha \rightsquigarrow \beta)$ e $p(\beta \mid \alpha)$.

Teorema 3.4.3 (Segundo resultado de trivialização de Lewis (1976, p. 302))**.** *Dado um conectivo $\rightsquigarrow$ (Definição 3.3.1) restrito a uma classe de funções de probabilidade fechada por condicionalização, não existem três proposições incompatíveis quando tomadas duas a duas que tenham todas probabilidades positivas.*

Demonstração. Suponha-se que o condicional $\rightsquigarrow$ está definido apenas para uma certa classe de funções de probabilidade fechada por condicionalização. Sejam p uma dessas funções e α, β duas proposições tais que $p(\alpha \wedge \beta) > 0$ e $p(\alpha \wedge \sim\beta) > 0$.
Então $p(\alpha \rightsquigarrow \beta) = p(\beta \mid \alpha) = p(\beta)$, pela Definição 3.3.1 e pelo Teorema 3.4.1.
Suponham-se haver, ainda, proposições γ, δ tais que β, γ, δ são incompatíveis quando tomadas duas a duas. Suponha-se, também, que $p(\gamma) > 0$ e $p(\delta) > 0$. Considere-se o caso em que $\alpha \leftrightarrow \beta \vee \gamma$. Temos que $p(\alpha) = p(\beta \vee \gamma)$ (por Equi). Então $p(\alpha \wedge \beta) = p([\beta \vee \gamma] \wedge \beta) = p(\beta)$,

[6] É nesse sentido que o condicional probabilístico $\rightsquigarrow$, se existisse, seria uma exceção a essa equivalência entre assertabilidade e probabilidade subjetiva: a assertabilidade de proposições da forma $\alpha \rightsquigarrow \beta$ seria igual não só a $p(\alpha \rightsquigarrow \beta)$, mas também a $p(\beta \mid \alpha)$.

dados os axiomas Equi e Disj. Mas observe-se que:

(3.189) $\varnothing \models \sim(\beta \wedge \gamma)$ hipótese

(3.190) $p(\alpha \wedge \beta) > 0$ hipótese

(3.191) $p(\gamma) > 0$ hipótese

(3.192) $p(\beta \mid \alpha) = \dfrac{p(\beta \wedge \alpha)}{p(\alpha)}$ Definição 3.0.2, (3.190)

Substituindo-se α por $\beta \vee \gamma$:

(3.193) $= \dfrac{p(\beta \wedge [\beta \vee \gamma])}{p(\beta \vee \gamma)}$ Definição de α, (3.190), (3.192)

(3.194) $= \dfrac{p(\beta)}{p(\beta \vee \gamma)}$ (3.193)

(3.195) $= \dfrac{p(\beta)}{p(\beta) + p(\gamma)}$ Disj, (3.189), (3.194)

(3.196) $< p(\beta)$ (3.190), (3.191), (3.195)

Ou seja, $p(\beta \mid \alpha) \neq p(\beta)$, o que contraria a igualdade obtida anteriormente pelo Teorema 3.4.1. Não podem existir, portanto, proposições β, γ, δ com valores de probabilidade positivos, mas incompatíveis quando tomadas duas a duas. ■

O segundo resultado de trivialização (Teorema 3.4.3) evidencia que o primeiro resultado (Teorema 3.4.1) não é uma exatamente uma contradição, uma vez que é possível que a linguagem do conectivo $\rightsquigarrow$ seja fraca a ponto de não conter três sentenças incompatíveis quando tomadas em pares, mas todas com probabilidades positivas. Por exemplo, pode haver uma linguagem em que toda sentença é ou tautológica ou contraditória (LEWIS, 1976, p. 300). É conveniente estipular a seguinte terminologia:

Definição 3.4.1 (Função de probabilidade trivial)**.** Uma função de probabilidade é *trivial* se nunca atribui valor positivo a mais de duas proposições incompatíveis entre si.

Então o primeiro resultado de trivialização (Teorema 3.4.1) implica o seguinte Corolário:

Corolário 3.4.4 (Lewis (1976, p. 300))**.** *Toda linguagem que contém um condicional probabilístico universal é trivial.*

Os dois primeiros resultados de Lewis (1976) já são destrutivos para a hipótese da Definição 3.3.1, portanto. Já o terceiro resultado concerne as consequências da existência de um condicional probabilístico para a classe das funções de probabilidade que correspondem a sistemas de crenças para os valores atribuídos por uma função de probabilidade definida para essa classe. Inicialmente, o Lema 3.4.5 evidencia que, se há pelo menos uma proposição de probabilidade intermediária, então condicionalizar sobre ela resulta em uma função de probabilidade "degenerada", isto é, que reduz-se a uma simples valoração (clássica).

Lema 3.4.5 (Degeneração da condicionalização sobre probabilidade intermediária)**.** *Se* $0 < p(\alpha) < 1$*, então* $p_\alpha(\beta) \in \{0,1\}$*; isto é, o contradomínio de* p_α *é o conjunto* $\{0,1\}$.

Demonstração.

(3.197)	$0 < p(\alpha) < 1$	hipótese
(3.198)	$0 < p(\sim\alpha) < 1$	Lema 3.1.3, (3.197)
(3.199)	$p_\alpha(\beta) = p(\alpha \rightsquigarrow \beta)$	Definição 3.0.3, (3.197)
(3.200)	$= p(\beta \mid \alpha)$	Definição 3.3.1, (3.199)
(3.201)	$= \dfrac{p(\alpha \wedge \beta)}{p(\alpha)}$	Definição 3.0.2, (3.200)

Caso I: $p(\alpha \wedge \beta) = 0$:

(3.202)	$p(\alpha \wedge \beta) = 0$	hipótese
(3.203)	$p_\alpha(\beta) = 0$	(3.201), (3.202)

Caso II: $p(\alpha \wedge \beta) > 0$:

(3.204)	$p(\alpha \wedge \beta) > 0$	hipótese
(3.205)	$p(\alpha \rightsquigarrow \beta) = p(\alpha \rightarrow \beta)$	Teorema 3.4.3, (3.204)
(3.206)	$p(\sim\alpha) = 0 \quad \text{ou} \quad p(\beta \mid \alpha) = 1$	Definição 3.3.1, Teorema 3.3.2, (3.205)
(3.207)	$p(\beta \mid \alpha) = 1$	(3.198), (3.206)
(3.208)	$p_\alpha(\beta) = 1$	(3.200), (3.207)

■

O resultado mais geral, no entanto, é que a existência de uma proposição que receba um valor de probabilidade intermediário de uma função p implica que qualquer outra proposição recebe de p um valor entre $0, p(\sim\alpha), p(\alpha),$ e 1.

Teorema 3.4.6 (Terceiro resultado de trivialização de Lewis (1976, p. 302))**.** *Se existe uma proposição* α *tal que* $0 < p(\alpha) < 1$*, então* $p(\beta) \in \{0, p(\sim\alpha), p(\alpha), 1\}$.

Demonstração. Seja α uma proposição tal que $0 < p(\alpha) < 1$. Então $0 < p(\sim\alpha) < 1$ (Lema 3.1.3). Pelo Lema 3.4.5, $p(\cdot \mid \alpha) \in \{0,1\}$ e $p(\cdot \mid \sim\alpha) \in \{0,1\}$. Mas $p(\beta) = p(\beta \mid \alpha)p(\alpha) + p(\beta \mid \sim\alpha)p(\sim\alpha)$ (Teorema 3.2.11), de modo que $p(\beta) \in \{0, p(\sim\alpha), p(\alpha), 1\}$. ■

Comentário. Note-se que $p(\beta) = 1$ ocorre apenas quando $p(\beta \mid \alpha) = p(\beta \mid \sim\alpha) = 1$, tal que $p(\beta) = p(\alpha) + p(\sim\alpha) = 1$. Trata-se de uma ocorrência do Lema 3.1.3, portanto.

Como existem funções de probabilidade subjetiva não-triviais, o condicional probabilístico $\rightsquigarrow$ hipotetizado na Definição 3.3.1 não pode vigorar nem mesmo apenas na classe que contém exatamente essas funções. Isso porque não pode ser interpretado de modo a garantir, para cada função de probabilidade da classe, que a probabilidade condicional seja igual à probabilidade do condicional:

> *Since some probability functions that represent possible systems of belief are not trivial, our indicative conditional is not a probability conditional for the class of all such probability functions. Whatever it may mean, it cannot possibly have a meaning such as to guarantee, for all possible subjective probability functions at once, that the probabilities of conditionals equal the corresponding conditional probabilities (LEWIS, 1976, p. 302–303).*

Lewis (1976) prossegue discutindo as alternativas para a formalização da assertabilidade de conteúdos semânticos condicionais, cotejando algumas propostas aventadas na literatura até aquela época. Como esta dissertação concentra-se sobre os resultados de trivialização em si, passa-se agora à análise de Fitelson (2015).

3.4.2 Os mais fortes: Fitelson (2015)

> All other (published) Lewisian triviality results are strictly weaker than ours, and there can be no stronger result along these lines [. . .].
>
> Fitelson (2015, p. 69).

Os resultados de trivialização de Fitelson (2015) são demonstrados por meio de um método algébrico, e essas demonstrações são verificadas por *prSAT*, uma biblioteca da suíte de Matemática computacional *Wolfram Mathematica*[7]. O método e sua implementação são expostos em Fitelson (2008). Serão comentados, a seguir, os aspectos desse método mais relevantes para a compreensão dos resultados de Fitelson (2015). Em particular, não se demonstrará sua validade; os interessados nessa demonstração são remetidos a Fitelson (2008).

O método de Fitelson (2008) baseia-se em formas normais disjuntivas (FNDs). Cabe lembrar que uma proposição está em FND se, e somente se, ela consiste de disjunções de *cláusulas*. Uma cláusula é uma conjunção de literais, e um *literal* é uma proposição que ou é atômica, ou é a negação de uma proposição atômica. A associatividade e a comutatividade da disjunção e da conjunção clássicas fazem com que seja razoável mencionar, por exemplo, *a conjunção*, no singular, de certos literais: embora várias conjunções correspondam aos mesmos literais, a equivalência semântica entre elas autoriza-nos a falar como se houvesse uma única conjunção possível (isto é, uma única cláusula). O mesmo vale, *mutatis mutandis*, para as disjunções de cláusulas[8]. O método algébrico baseia-se na possibilidade de reescrever qualquer proposição de LPC como uma proposição em FND[9]. Dito de outro modo, baseia-se no fato de que, para toda proposição clássica, existe uma proposição em FND que lhe é equivalente. Outro fato conhecido a respeito de lógicas clássicas é que todas as valorações possíveis para uma

[7] Disponível em `<http://fitelson.org/PrSAT>` (acesso em 7/10/2019).

[8] Para os propósitos deste trabalho, poder-se-ia até dizer que uma cláusula é não uma conjunção de literais, mas a classe de equivalência dessas conjunções, e que uma FND não é uma disjunção de cláusulas, mas a respectiva classe de equivalência.

[9] Essa possibilidade não será demonstrada aqui, sendo um teorema bastante conhecido de lógica clássica. A maioria dos manuais de Lógica dedica uma seção a formas normais.

determinada proposição α podem ser visualizadas em uma tabela de 2^n linhas, se n é o número de variáveis proposicionais distintas que ocorrem em α[10]. Essas tabelas são conhecidas por "tabelas de verdade". Convencione-se que cada cláusula seja denotada por $\wedge_i$, em que $1 \leq i \leq 2^n$ é o índice da respectiva linha da tabela de verdade. Dada uma proposição α qualquer, pode-se encontrar uma FND equivalente a α executando-se o seguinte algoritmo:

1. Escrever cada uma das n variáveis proposicionais distintas que ocorrem em α em uma coluna diferente de uma tabela de verdade;
2. Preencher cada célula da tabela de modo a representar todas as 2^n combinações possíveis de `t` e `f` para as n variáveis proposicionais;
3. Em cada célula de uma nova coluna, escrever a cláusula $\wedge_i$ em que cada variável proposicional daquela linha aparece afirmada caso seja verdadeira (naquela linha), e negada em caso contrário.
4. Preencher cada célula de uma nova coluna com o valor de verdade de α naquela valoração;
5. Definir o conjunto $\mathrm{I}_\alpha := \{1 \leq i \leq 2^n : \alpha$ é verdadeira na i-ésima linha da tabela de verdade$\}$.

Então α será equivalente à seguinte proposição em FND:

$$\alpha \leftrightarrow \begin{cases} \perp & \text{se } \mathrm{I}_\alpha = \varnothing \\ \bigvee\limits_{i \in \mathrm{I}_\alpha} \wedge_i & \text{do contrário} \end{cases}$$

Exemplo. Seja $\mathrm{L}_{a,b}$ a lógica proposicional clássica cuja linguagem contém exatamente duas variáveis proposicionais, a e b. Então seja α a proposição $(a \rightarrow b) \wedge (a \vee \sim b)$. Executar os passos descritos anteriormente resulta na seguinte tabela de verdade:

	a	b	α	$\wedge_i$
1	t	t	t	$a \wedge b$
2	t	f	f	$a \wedge \sim b$
3	f	t	f	$\sim a \wedge b$
4	f	f	t	$\sim a \wedge \sim b$

α é verdadeira nas linhas 1 e 4, de modo que é equivalente à proposição em FND $(a \wedge b) \vee (\sim a \vee \sim b)$, porque essa é a disjunção das cláusulas $\wedge_1$ e $\wedge_4$.

O método algébrico de Fitelson (2008) começa pela montagem de uma tabela semelhante a essas, denominada "tabela estocástica" (FITELSON, 2008, p. 112). Trata-se de uma tabela de verdade acrescida de uma coluna ocupada pelos valores de probabilidade que a função p atribui a cada cláusula. Por exemplo, uma linguagem proposicional clássica que contenha exatamente três variáveis proposicionais (a, b, c) corresponderá à seguinte tabela estocástica:

[10] O valor da base dessa exponenciação é o número de valores de verdade da lógica em questão. As tabelas de uma lógica trivalorada, por exemplo, contêm sempre 3^n linhas. Muitas lógicas não-clássicas também são capturáveis por matrizes finitas (isto é, "tabeláveis"), mas não todas.

	a	b	c	$\wedge_i$	$p(\cdot)$
1	t	t	t	$a \wedge b \wedge c$	p_1
2	t	t	f	$a \wedge b \wedge \sim c$	p_2
3	t	f	t	$a \wedge \sim b \wedge c$	p_3
4	t	f	f	$a \wedge \sim b \wedge \sim c$	p_4
5	f	t	t	$\sim a \wedge b \wedge c$	p_5
6	f	t	f	$\sim a \wedge b \wedge \sim c$	p_6
7	f	f	t	$\sim a \wedge \sim b \wedge c$	p_7
8	f	f	f	$\sim a \wedge \sim b \wedge \sim c$	p_8

Tabela 5 – A tabela estocástica para uma lógica proposicional clássica com exatamente três variáveis proposicionais (FITELSON, 2008, p. 112).

A chave do método de Fitelson (2008) é que, assim como, uma proposição é verdadeira exatamente quando uma das respectivas cláusulas o é, também seu valor de probabilidade é a soma dos valores de probabilidade dessas cláusulas. E o valor atribuído por uma função de probabilidade condicional da forma $p(\alpha \mid \beta)$ é a razão entre a soma dos valores de probabilidade das cláusulas nas quais $\alpha \wedge \beta$ é verdadeira e a soma dos valores de probabilidade das cláusulas nas quais β é verdadeira. Ou seja, existe um isomorfismo entre a relação lógica de disjunção e a operação algébrica de soma: "this gives us a faithful translation from the language [. . .] into the language of simple real algebra (*viz.*, high-school algebra)" (FITELSON, 2008, p. 112). Por exemplo, na Tabela 5, $p(b) = p_1 + p_2 + p_5 + p_6$, porque $b \leftrightarrow (\wedge_1 \vee \wedge_2 \vee \wedge_5 \vee \wedge_6)$ (FITELSON, 2008, p. 112). E, ainda, $p(b \wedge c) = p_1 + p_5$, porque b e c são simultaneamente verdadeiras exatamente nas cláusulas $\wedge_1$ e $\wedge_5$. Já $p(b \mid c) = \frac{p_1+p_5}{p_1+p_3+p_5+p_7}$. Esse isomorfismo não deve surpreender, considerando-se o teorema de representação de Stone e que a probabilidade de Kolmogorov definida sobre proposições é formalmente equivalente àquela definida sobre conjuntos, como já comentado.

O isomorfismo entre as operações algébricas das tabelas estocásticas e as relações lógicas clássicas é garantido, explica Fitelson (2008, p. 113), caso sejam impostas duas restrições algébricas aos valores $p_1, \ldots, p_{2^n}$ de uma tabela estocástica qualquer com n variáveis proposicionais distintas:

i) para todo $1 \leq i \leq 2^n$, $p_i \geq 0$; e

ii) $\sum_{i=1}^{2^n} p_i = 1$.

Como a probabilidade condicional (binária) é definida em termos da probabilidade absoluta, essas duas condições também bastam para que as funções de probabilidade condicional adéquem-se ao sistema de Kolmogorov (FITELSON, 2008, p. 113). Prova-se então o seguinte teorema:

Teorema 3.4.7 (Fitelson (2015, p. 69))**.** *Dada a definição do conectivo* $\rightsquigarrow$ *(Definição 3.3.1), o lema da exportação/importação (Lema 3.3.5) equivale ao lema da probabilidade condicional composta (Lema 3.3.3): se* $p(\alpha \wedge \gamma) > 0$*, é o caso que* $p(\alpha \rightsquigarrow [\beta \rightsquigarrow \gamma]) = p((\alpha \wedge \beta) \rightsquigarrow \gamma) = p((\beta \wedge \alpha) \rightsquigarrow \gamma)$ *se, e somente se, também é o caso que* $p(\alpha \rightsquigarrow \beta \mid \gamma) = p(\beta \mid \alpha \wedge \gamma)$*.*

Demonstração.

Direção ($\Rightarrow$):

(3.209)	$p(\alpha \rightsquigarrow (\beta \rightsquigarrow \gamma)) = p((\alpha \wedge \beta) \rightsquigarrow \gamma)$	hipótese (Lema 3.3.5)
(3.210)	$= p(\gamma \mid \alpha \wedge \beta)$	hipótese (Definição 3.3.1), (3.209)
(3.211)	$= p(\gamma \mid \beta \wedge \alpha)$	Teorema 2.1.1, Lema 3.1.6, (3.210)
(3.212)	$p(\alpha \rightsquigarrow (\beta \rightsquigarrow \gamma)) = p(\beta \rightsquigarrow \gamma \mid \alpha)$	hipótese (Definição 3.3.1)
(3.213)	$p(\beta \rightsquigarrow \gamma \mid \alpha) = p(\gamma \mid \beta \wedge \alpha)$	(3.211), (3.212)
(3.214)	$p(\alpha \rightsquigarrow \beta \mid \gamma) = p(\beta \mid \alpha \wedge \gamma)$	Substituição de β por α, γ por β, e α por γ, (3.213)

Direção ($\Leftarrow$):

(3.215)	$p(\alpha \rightsquigarrow \beta \mid \gamma) = p(\beta \mid \alpha \wedge \gamma)$	hipótese (Lema 3.3.3)
(3.216)	$p(\alpha \rightsquigarrow (\beta \rightsquigarrow \gamma)) = p((\beta \rightsquigarrow \gamma) \mid \alpha)$	hipótese (Definição 3.3.1)
(3.217)	$= p(\gamma \mid \beta \wedge \alpha)$	(3.215), (3.216)
(3.218)	$= p(\gamma \mid \alpha \wedge \beta)$	Teorema 2.1.1, Lema 3.1.6, (3.217)
(3.219)	$= p((\alpha \wedge \beta) \rightsquigarrow \gamma)$	hipótese (Definição 3.3.1), (3.218)

■

$\rightsquigarrow$ não é um conectivo verofuncional, de modo que, na tabela estocástica de α, β e $\alpha \rightsquigarrow \beta$, sendo α e β duas proposições quaisquer, a proposição condicional comporta-se como uma terceira proposição: "[. . .] I will be assuming (without loss of generality) that α, β, and $\alpha \rightsquigarrow \beta$ are logically independent of each other. If there were logical dependencies between them, then this would only serve to strengthen our triviality result" (FITELSON, 2015, p. 73). É esta, portanto, a tabela estocástica dessas proposições:

	α	β	$\alpha \rightsquigarrow \beta$	p
1	t	t	t	a
2	t	t	f	b
3	t	f	t	c
4	t	f	f	d
5	f	t	t	e
6	f	t	f	f
7	f	f	t	g
8	f	f	f	h

Tabela 6 – Os valores de verdade e as probabilidades de duas proposições quaisquer e de uma proposição em que elas são antecedente e consequente de um condicional probabilístico $\rightsquigarrow$ (FITELSON, 2015, p. 70).

Como se vê, denominaram-se “a”, “b”, . . . , “h” os valores que p atribui a cada valoração. A seguir, alguns lemas necessários para o teorema de trivialização de Fitelson (2015).

Lema 3.4.8 (Instância $\gamma = {\sim}\beta$ do Lema 3.3.3)**.** *Se* $p(\alpha \wedge {\sim}\beta) > 0$*, então* $c = g = 0$.

Demonstração.

(3.220) $p(\alpha \wedge {\sim}\beta) > 0$ hipótese

(3.221) $p(\alpha \rightsquigarrow \beta \mid \gamma) = p(\beta \mid \alpha \wedge \gamma)$ Lema 3.3.3

Seja $\gamma = {\sim}\beta$:

(3.222) $p(\alpha \rightsquigarrow \beta \mid {\sim}\beta) = p(\beta \mid \alpha \wedge {\sim}\beta)$ (3.220), (3.221)

(3.223) $= 0$ Lema 3.2.10, (3.222)

(3.224) $p(\alpha \rightsquigarrow \beta \mid {\sim}\beta) = \dfrac{p([\alpha \rightsquigarrow \beta] \wedge {\sim}\beta)}{p({\sim}\beta)}$ Definição 3.0.2, (3.221)

(3.225) $p(\alpha \rightsquigarrow \beta \mid {\sim}\beta) = \dfrac{c + g}{c + d + g + h}$ Tabela 6, (3.224)

(3.226) $0 = \dfrac{c + g}{c + d + g + h}$ (3.223), (3.225)

(3.227) $0 = c + g$ (3.226)

(3.228) $c = g = 0$ Cdmn, (3.227)

■

Lema 3.4.9 (Instância $\gamma = \alpha \rightarrow \beta$ do Lema 3.3.3)**.** *Se* $p(\alpha \wedge \beta) > 0$*, então* $b = f = h = 0$.

Demonstração.

(3.229) $p(\alpha \wedge \beta) > 0$ hipótese

(3.230) $p(\alpha \rightsquigarrow \beta \mid \gamma) = p(\beta \mid \alpha \wedge \gamma)$ Lema 3.3.3

Seja $\gamma = \alpha \rightarrow \beta$ e considerando-se que $(\alpha \wedge \beta) \leftrightarrow (\alpha \wedge [\alpha \rightarrow \beta])$ (Teorema 2.1.7):

(3.231) $p(\alpha \rightsquigarrow \beta \mid \alpha \rightarrow \beta) = p(\beta \mid \alpha \wedge [\alpha \rightarrow \beta])$ (3.229), (3.230)

(3.232) $= p(\beta \mid \alpha \wedge \beta)$ Teorema 2.1.7, Lema 3.1.6, (3.231)

(3.233) $= 1$ Lema 3.2.1, (3.232)

(3.234) $p(\alpha \rightsquigarrow \beta \mid \alpha \rightarrow \beta) = \dfrac{p([\alpha \rightsquigarrow \beta] \wedge [\alpha \rightarrow \beta])}{p(\alpha \rightarrow \beta)}$ Definição 3.0.2

O valor de g pode ser desconsiderado, conforme o último Lema. Então:

(3.235) $= \dfrac{a + e}{a + b + e + f + h}$ Tabela 6, (3.230)

(3.236) $\dfrac{a + e}{a + b + e + f + h} = 1$ (3.233), (3.235)

(3.237) $b + f + h = 0$ (3.236)

(3.238) $b = f = h = 0$ Cdmn, (3.237)

■

É importante lembrar que $a + \ldots + h = 1$, devido às restrições às quais as tabelas estocásticas devem se conformar. Como $b = c = f = g = h = 0$, $a + d + e = 1$. Além disso, as premissas dos Lemas 3.4.8 e 3.4.9 mantêm-se; a do primeiro será necessária novamente: $p(\alpha \wedge \sim\beta) > 0$. Isso significa que $d > 0$, porque, dentre as cláusulas da Tabela 6 que ainda podem ter probabilidade positiva, d é a unica em que ocorrem α e $\sim\beta$.

Lema 3.4.10 (Instância $\gamma = \top$ do Lema 3.3.3)**.** $e = 0,\ d = 1 - a$ e $a, d \in (0, 1)$.

Demonstração.

(3.239) $a + d + e = 1$ Tabela 6, Lemas 3.4.8 e 3.4.9

(3.240) $d > 0$ Hipótese do Lema 3.4.8

(3.241) $p(\alpha \rightsquigarrow \beta \mid \gamma) = p(\beta \mid \alpha \wedge \gamma)$ Lema 3.3.3

Seja $\gamma = \top$:

(3.242) $p(\alpha \rightsquigarrow \beta) = p(\beta \mid \alpha \wedge \top)$ Lema 3.1.7, (3.241)

(3.243) $= p(\beta \mid \alpha)$ Lema 3.1.6, (3.242)

Os valores de b, c, f, g, h podem ser desconsiderados, conforme os últimos Lemas. Então:

(3.244) $a + e = \dfrac{a}{a + d}$ Tabela 6, (3.243)

(3.245) $a^2 + (d + e - 1)a + de = 0$ (3.244)

Dada a premissa (3.239), pode-se substituir $d + e$ por $1 - a$:

(3.246) $a^2 + (1 - a - 1)a + de = 0$ (3.239), (3.245)

(3.247) $de = 0$ (3.246)

(3.248) $e = 0$ (3.240), (3.247)

(3.249) $d = 1 - a$ (3.239), (3.248)

Esse resultado é confirmado pela resolução da equação de segundo grau $a^2 + (d-1)a = 0$: suas raízes são 0 e $1 - d$; mas sabe-se que $a > 0$. ■

Note-se que a instância $\gamma = \top$ do Lema 3.3.3 nada mais é que a Definição 3.3.1, como fica claro em sua demonstração. Os Lemas 3.4.8–3.4.10 levam à seguinte atualização da Tabela 6:

α	β	$\alpha \rightsquigarrow \beta$	p
t	t	t	a
t	t	f	0
t	f	t	0
t	f	f	$1 - a$
f	t	t	0
f	t	f	0
f	f	t	0
f	f	f	0

Tabela 7 – Revisão da Tabela 6 conforme os Lemas 3.4.8–3.4.10 (FITELSON, 2015, p. 72).

Essa tabela evidencia o seguinte Corolário:

Corolário 3.4.11. *Os Lemas 3.4.8–3.4.10 implicam que as únicas cláusulas (linhas) da Tabela 6 às quais a função de probabilidade atribui valor positivo são* $p(\alpha \wedge \beta \wedge \alpha \rightsquigarrow \beta)$ *e* $p(\alpha \wedge {\sim}\beta \wedge {\sim}(\alpha \rightsquigarrow \beta))$.

Pode-se, enfim, provar o seguinte resultado:

Teorema 3.4.12 (Trivialização de Fitelson (2015))**.** *Se* $p(\alpha \wedge \beta) > 1$ *e* $p(\alpha \wedge {\sim}\beta) > 1$*, então o Lema 3.3.3 equivale a* $p(\alpha \wedge (\beta \leftrightarrow [\alpha \rightsquigarrow \beta])) = 1$.

Demonstração.

(3.250) $a + (1 - a) = 1$ aritmética

(3.251) $p(\alpha \wedge \beta \wedge \alpha \rightsquigarrow \beta) + p[\alpha \wedge {\sim}\beta \wedge {\sim}(\alpha \rightsquigarrow \beta)] = 1$ Tabela 7, (3.250)

(3.252) $p([\alpha \wedge \beta \wedge \alpha \rightsquigarrow \beta] \vee [\alpha \wedge {\sim}\beta \wedge {\sim}(\alpha \rightsquigarrow \beta)]) = 1$ Teorema 2.1.10, Disj, (3.251)

(3.253) $p(\alpha \wedge (\beta \leftrightarrow [\alpha \rightsquigarrow \beta])) = 1$ Teorema 2.1.10, Equi, (3.252)

■

O Teorema 3.4.12 é muito forte, como comenta Fitelson (2015, p. 72); isso fica explícito na prova do seguinte Corolário:

Corolário 3.4.13. *Se* $p(\alpha \wedge (\beta \leftrightarrow [\alpha \rightsquigarrow \beta])) = 1$*, então* β *e* α *são proposições independentes.*

Demonstração.

(3.254)	$p(\alpha \wedge \beta \leftrightarrow [\alpha \rightsquigarrow \beta]) = 1$	hipótese
(3.255)	$p(\beta \leftrightarrow [\alpha \rightsquigarrow \beta]) = 1$	Lema 3.2.9, (3.254)
(3.256)	$p(\beta) = p(\alpha \rightsquigarrow \beta)$	Lema 3.2.4, (3.255)
(3.257)	$= p(\beta \mid \alpha)$	Definição 3.3.1, (3.256)

■

Em resumo, procurou-se dissecar as demonstrações das chamadas "leis do excesso" e dos resultados de trivialização conforme expostos por Lewis (1976) e Fitelson (2015). Em ambos os textos, como de praxe em artigos acadêmicos, as provas são bastantes sucintas e por vezes é difícil entender sobre quais premissas exatamente repousam. Espera-se que a análise deste capítulo, embora possivelmente maçante, tenha lançado alguma luz sobre a compreensão desses resultados. A pergunta que norteou esta pesquisa foi: "o que aconteceria com esses teoremas caso a lógica subjacente fosse paraconsistente"? O capítulo 5 pretende retornar o que foi estudado até aqui para hipotetizar respostas. Antes, no entanto, o capítulo 4 introduz rapidamente a paraconsistência.

4 Controlando explosões: sobre paraconsistência

> We are the music makers,
> And we are the dreamers of dreams,
> Wandering by lone sea-breakers,
> And sitting by desolate streams;—
> World-losers and world-forsakers,
> On whom the pale moon gleams:
> Yet we are the movers and shakers
> Of the world for ever, it seems.
>
> Arthur O'Shaughnessy, *Ode*.

Os conceitos de inconsistência e trivialidade são equivalentes e intersubstituíveis na lógica clássica. É o que mostra o chamado princípio da explosão, também conhecido como *ex falso quodlibet* (da falsidade, tudo (se segue)):

Lema 4.0.1 (Princípio da explosão)**.** $\{\alpha, \sim\alpha\} \models_{LPC} \beta$.

Demonstração. Considere-se um sistema clássico com *modus ponens* como a única regra de inferência. Sejam α uma proposição tanto verdadeira quanto falsa e β uma proposição arbitrária. $\alpha \rightarrow \beta$ é verdadeira, dados a tabela de verdade da implicação material e que $\sim\alpha$. Mas então, como $\alpha \rightarrow \beta$ e α são ambas verdadeiras, β também o é, por *modus ponens*. ■

Assim, se um sistema clássico é inconsistente, também é trivial. As lógicas paraconsistentes são justamente as que distinguem entre inconsistência e trivialidade, de modo que, se L é uma lógica paraconsistente, então existem $\alpha, \beta \in L$ tais que $\{\alpha, \neg\alpha\} \not\models_{L} \beta$.

Uma ideia intuitiva quanto à definição de paraconsistência é a de que as lógicas paraconsistentes rejeitam o esquema de axioma clássico $\sim(\alpha \wedge \sim\alpha)$. Isso não se verifica, no entanto. Muitas lógicas paraconsistentes mantêm esse esquema, e esse é o caso inclusive de LFI1, o sistema que particularmente enfocado nesta dissertação (MARCOS, 2005, p. XXXV).

O surgimento das lógicas paraconsistentes insere-se, é claro, no contexto maior das primeiras lógicas não-clássicas. Como já comentado, a passagem do século XIX para o XX foi marcada por uma crise nos fundamentos da Matemática que fomentou um desenvolvimento da Lógica sob a égide dos projetos formalista, intuicionista e logicista. Isso consolidou a Matemática, ou o fazer matemático, como o escopo da Lógica, mas também

> contribuiu, sobretudo, para depurar e analisar com acurácia os fundamentos da logicidade clássica, seja pelo estabelecimento claro e formalização de seus axiomas e teoremas, seja pela análise atenta de metarresultados importantes

> como completude, consistência e decidibilidade desses sistemas lógicos e pelo aprimoramento dos métodos formais aí empregados (GOMES, 2013, p. 281).

Gomes (2013, p. 281–282) atribui o surgimento das lógicas não-clássicas a essa depuração promovida pela crise nos fundamentos da Matemática e também a uma reavaliação da lógica aristotélica naquele momento: "um intenso debate teórico seguiu-se aí e em seu bojo, sugeriu-se o refinamento e a releitura das premissas hermenêuticas tradicionais, de acordo com a qual Aristóteles teria apenas sugerido e dado suporte ao paradigma lógico clássico". Um exemplo é o trabalho de Jan Łukasiewicz (1878–1956) de 1951 sobre a silogística aristotélica na óptica da lógica contemporânea (GOMES, 2013, p. 285). Para entender o surgimento das lógicas ditas heterodoxas, convém observar que naquela época surgiam as geometrias não-euclidianas, em boa parte responsáveis pelo espírito de exploração formal da noção de consequência, de *entailment*, tão determinante para as primeiras lógicas não-clássicas:

> Uma vez que há [geometrias] alternativas, e que todas merecem estudo, não compete ao matemático afirmar os axiomas que tornam uma das [geometrias] alternativas naquilo que ela é. A sua função é dizer o que é que se segue logicamente de um dado conjunto de axiomas. O fato de se considerar a geometria tradicional desta maneira conduziu à necessidade da formulação explícita de todos os seus pressupostos, a fim de se cumprir rigorosamente o programa que atribuímos a Euclides; *e.g.* os *Grundlagen der Geometrie* de Hilbert de 1899 (Kneale e Kneale (1991), p. 389, *apud* (GOMES, 2013, p. 283)).

A íntima relação entre os surgimentos das geometrias não-euclidianas, por um lado, e das lógicas não-clássicas, por outro, foi consciente, e não uma aproximação só reconhecida em retrospecto. Isto é, o desenvolvimento da Geometria tornou-a um modelo para a Lógica (GOMES, 2013, p. 284, 289).

Embora seja sempre difícil distinguir entre precursores e fundadores e apontar pioneiros em uma teoria ou linha de pesquisa, pode-se considerar que as lógicas não-clássicas surgiram por volta de 1910, com os trabalhos de Łukasiewicz e Nicolai A. Vasiliev (1880–1940). Łukasiewicz, por exemplo, ocupava-se à época com o princípio de não-contradição em Aristóteles. Bem mais que comentar o filósofo antigo, a pesquisa era um estudo do princípio em si, incluindo aí a possibilidade de não se tratar de uma lei, no sentido forte que a palavra tem, do pensamento.

Já as primeiras (e talvez, até hoje, mais influentes) lógicas paraconsistentes foram os sistemas C_n, $0 < n < \omega$, de Newton C. A. da Costa (1929–), publicadas em 1974. Neles, a trivialidade diferencia-se da contraditoriedade (e esta equivale à inconsistência), devido à adoção de três princípios:

C1 em geral, as fbfs não satisfazem o princípio da explosão, mas o satisfaz toda fórmula α tal que $\neg(\alpha \wedge \neg\alpha)$ é o caso; e

C2 se α *e* β satisfazem o princípio da explosão, então também satisfazem-no $\alpha \wedge \beta$, $\alpha \vee \beta$, e $\alpha \to \beta$.

Os sistemas C_n^+, $0 < n < \omega$, introduzidos pelo mesmo autor em 1995, são aqueles em que o princípio C2 é substituído por C2':

C2' se α *ou* β satisfazem o princípio da explosão, então também satisfazem-no $\alpha \wedge \beta$, $\alpha \vee \beta$, e $\alpha \rightarrow \beta$.

Entre as primeiras lógicas paraconsistentes multivaloradas, destacam-se P^1 (1973) e J3 (Ítala M. L. D'Ottaviano (1944–), da Costa, 1970).

Qual a motivação para as lógicas paraconsistentes? Essa pergunta é parte de uma questão maior, a de qual é a necessidade ou significado, se é que existem, de lógicas não-clássicas em geral; por que criar alternativas aos sistemas clássicos? Uma resposta é a de que não há razão para tanto, porque a lógica clássica representa (ou normatiza, o que é ainda outra problemática) a totalidade do raciocínio dedutivo. Argumentar-se-á, aqui, em favor do contrário.

Dentre as várias justificativas que poderiam ser dadas para a importância das lógicas paraconsistentes (e, em geral, das não-clássicas), a que será enfocada aqui é a ideia de que o tratamento de informação, uma das áreas de aplicação da Lógica mais importantes atualmente, é um dos campos para os quais as lógicas paraconsistentes são particularmente adequadas. Isto é, a lógica paraconsistente a ser adotada aqui deverá destacar-se por sua efetividade em manejar informações. Convém ainda que ela seja *maximal* em relação a LPC, condição que os sistemas C_n, $0 < n < \omega$, e C_n^+, $0 < n < \omega$, não satisfazem.

Definição 4.0.1 (Maximalidade (CARNIELLI; MARCOS; DE AMO, 2004, p. 135))**.** Uma lógica L_1 é *maximal em relação a* uma lógica L_2 se ambas compartilham a mesma linguagem, demonstra-se todo teorema de L_1 em L_2, e, dado um teorema de L_2 que não é teorema de L_1, adicioná-lo a L_1 como novo esquema de axioma basta para demonstrar todos os teoremas de L_2.

Pode-se pensar que uma lógica maximal em relação a uma outra "é muito semelhante" a esta. Torna-se natural, assim, optar por LFI1: é um sistema concebido especificamente para o tratamento de informação, e sua maximalidade em relação a LPC evita que a literatura prévia sobre bases de dados seja desperdiçada (CARNIELLI; MARCOS; DE AMO, 2004, p. 120). Conforme será visto adiante, embora LFI1 seja formalmente equivalente a J3, esta última é uma lógica menos apta ao manejo de informação.

4.1 Justificando as lógicas paraconsistentes: o caso das bases de dados relacionais

> To accept everything is an exercise, to understand everything a strain.
>
> G. K. Chesterton, *Orthodoxy*.

Um contexto que evidencia a relevância da paraconsistência é o do conhecimento empírico, tanto científico quanto cotidiano, contraposto ao formal, como o matemático. Um argumento clássico é válido se, e somente se, sempre que suas premissas são todas verdadeiras, sua conclusão também o é. Não é assim, contudo, que se raciocina com dados empíricos: nesse âmbito, raramente é razoável dizer que uma sentença é absolutamente verdadeira, ou falsa. E ainda que se comece com sentenças todas muito prováveis, pode ocorrer que elas acarretem uma conclusão improvável[1]. Além disso, o conhecimento formal é atemporal: nunca é necessário revisar o que já se conhece quando um novo conhecimento é adquirido. A contraparte lógica da atemporalidade é a monotonicidade, que, como visto anteriormente, é condição necessária para a existência de uma relação de dedutibilidade semântica em uma lógica (clássica). Isto é, inferir uma proposição é irrevogável. A monotonicidade parece inadequada ao conhecimento empírico porque agentes racionais estão constantemente ajustando suas crenças diante de novas evidências. Na Ciência, por exemplo, cada novo fato é pesado contra o corpo de evidências já existente e as teorias associadas a ele, de modo a somar-se a essas evidências ou promover a revisão desse corpo. No cotidiano, são comuns casos como o de uma pessoa que acredita estar sendo pontual em um compromisso, mas então descobre que seu relógio esteve parado, o que significa que a crença sobre pontualidade deve ser retratada. De fato, esse ajuste contínuo parece ser inerente à parte da racionalidade que se ocupa dos fenômenos empíricos.

Uma circunstância que evidencia algumas dessas particularidades do conhecimento empírico é o armazenamento de informações em bancos de dados, como o cadastro de dados pessoais na Receita Federal. Comentar-se-á, brevemente, como lógicas paraconsistentes podem ser mais adequadas que sistemas clássicos nesses contextos. Uma *base de dados relacional* é uma coleção finita de relações finitas que codificam informações. Essas informações são manipuladas e consultadas, geralmente, por meio de linguagens de programação conhecidas como "linguagens de consulta"; uma das mais populares é a linguagem de consulta estruturada (SQL). Cada modificação em uma base está condicionada à satisfação das suas *condições de integridade*, expressas por sentenças de uma lógica de primeira ordem (CARNIELLI; MARCOS; DE AMO, 2004, p. 116). Isto é, cada inserção, remoção ou edição de informação só é efetuada se não resultar na infração de alguma exigência daquela base. Por exemplo, um cadastro de pessoas físicas em um sistema poderia requerer que cada indivíduo tenha exatamente uma data de nascimento. Essa expectativa poderia ser representada pela seguinte sentença[2]:

$$\forall x\Big(\mathrm{P}x \to \big(\exists y(\mathrm{D}y \land \mathrm{N}xy \land \forall z(\mathrm{D}z \land \mathrm{N}xz \to z = y))\big)\Big)$$

Como seria de se esperar, uma condição de integridade comum é a de consistência. Manter as bases de dados consistentes tornou-se uma tarefa muito complexa com o desenvolvimento da tecnologia, no entanto, porque tornou-se possível atualizá-las a partir de várias fontes, por exemplo. Embora cada base de dados local esteja sempre consistente, garantir que suas várias

[1] Para um exemplo, conferir o chamado "paradoxo do prefácio".
[2] Considere-se que $\mathrm{P}x$ significa "x é uma pessoa", $\mathrm{D}x$ denota "x é uma data", e $\mathrm{N}xy$ expressa "x nasceu em y".

versões, hospedadas em outras máquinas, também não se contradizem vem tornando-se uma tarefa cada vez mais custosa (em termos de recursos computacionais) (CARNIELLI; MARCOS; DE AMO, 2004, p. 116). Outra fonte de inconsistências em bases de dados relacionais é a modificação das próprias condições de integridade por seus usuários. Bases de dados que admitem essa possibilidade são denominadas “evolucionárias” (CARNIELLI; MARCOS; DE AMO, 2004, p. 117). O formato tradicional, em que as condições de integridade são estipuladas durante a criação da base e não podem ser alteradas, tem a vantagem de ser mais “seguro” (isto é, menos propenso à inconsistência), mas também pode ser considerado restritivo demais.

Isto é, como todo procedimento operado por humanos, o cadastro de informações em bancos de dados está sujeito a erros que, acumulados, resultam por vezes em um corpo de informações inconsistente. Seria irracional, no entanto, inferir daí qualquer afirmação, e é por isso que lógicas paraconsistentes podem ser mais adequadas ao manejo de informações em bancos de dados que sistemas clássicos.

4.2 As lógicas da inconsistência formal

> [. . .] our approach regards contradictions and inconsistencies as phenomena waiting to be formally treated.
>
> Carnielli, Marcos e de Amo (2004, p. 120).

As lógicas da inconsistência formal são uma fração grande e bastante importante das lógicas paraconsistentes. Diferentemente de sistemas como J3, essas lógicas encaram as inconsistências como “fenômenos esperando serem formalmente tratados” (vide citação na epígrafe desta seção). A principal peculiaridade delas é internalizarem o conceito de consistência definindo um operador de consistência $\circ$ e um operador de inconsistência $\bullet$. Geralmente, um é dado como conectivo primitivo, e o outro, definido a partir dele, mas não necessariamente é assim.

A lógica J3 é um sistema modal paraconsistente trivalorado, cuja motivação era prover uma alternativa à lógica discussiva de Stanisław Jaśkowski (1906–1965), que foi, por sua vez, uma precursora das lógicas paraconsistentes. A assinatura de J3 é $\Sigma_{J3} = \{\neg, \nabla, \vee\}$, e a semântica desses conectivos é dada pelas tabelas da Figura 5. Suas valorações tem como contradomínio o conjunto $\{0, {}^1\!/\!_2, 1\}$, em que 1 e ${}^1\!/\!_2$ são os valores distinguidos.

α	$\neg\alpha$
1	0
½	½
0	1

α	$\nabla\alpha$
1	1
½	1
0	0

α	β	$\alpha \vee \beta$
1	1	1
1	½	1
1	0	1
½	1	1
½	½	½
½	0	½
0	1	1
0	½	½
0	0	0

Figura 5 – As tabelas de verdade dos conectivos de J3.

Podem ser facilmente definidos conectivos adicionais de significado intuitivo, a saber, uma negação forte, que emula a clássica, e um operador de consistência, que retorna 1 para valores de verdade extremos[3]:

α	$\sim\alpha$
1	0
½	0
0	1

α	$\circ\alpha$
1	1
½	0
0	1

Figura 6 – As tabelas de verdade da negação forte e da consistência em J3.

J3 reapareceu várias vezes na literatura, construída de diversas formas e por diferentes motivos (CARNIELLI; CONIGLIO, 2016, p. 141–142). Em 2004, Batens e De Clerq introduziram uma lógica de primeira ordem chamada CLuNs, cujo fragmento proposicional é Φ_v, que, por sua vez, é completo (e correto) em relação a J3. Coniglio e Silvestrini generalizaram a noção de quasi-verdade de Mikenberg, da Costa e Chuaqui com uma lógica de primeira ordem chamada LPT1, cujo fragmento proposicional, LPT, coincide com J3. Recentemente, em 2015, Löwe e Tarafder utilizaram uma versão de J3 denominada $\mathbb{PS}_3$ para construir uma teoria de conjuntos paraconsistente.

Como mencionado antes, a lógica LFI1 também é formalmente equivalente a J3, isto é, os conectivos primitivos de ambas são interdefiníveis, de modo que existem funções que mapeiam proposições de um sistema no outro. A diferença entre as duas lógicas é que o fato de J3 priorizar a noção de possibilidade (∇) e LFI1 enfatizar a noção de consistência ($\circ$) torna este último sistema mais apto ao manejo de bancos de dados (CARNIELLI; MARCOS; DE AMO, 2004, p. 120). LFI1 foi introduzida em de Amo, Carnielli e Marcos (2002) e sua assinatura é o

3 Com efeito, $\sim\alpha := \neg\nabla\alpha$ e $\circ\alpha := \neg\nabla\alpha \vee \neg\nabla\neg\alpha$.

conjunto $\Sigma_\bullet = \{\neg, \bullet, \wedge, \vee, \rightarrow\}$[4]. A semântica dos conectivos é dada pelas tabelas da Figura 7, e, como em J3, suas valorações têm como contradomínio o conjunto $\{0, 1/2, 1\}$, sendo 1 e $1/2$ os valores distinguidos.

α	$\neg\alpha$
1	0
$1/2$	$1/2$
0	1

α	$\bullet\alpha$
1	0
$1/2$	1
0	0

α	β	$\alpha \wedge \beta$
1	1	1
1	$1/2$	$1/2$
1	0	0
$1/2$	1	$1/2$
$1/2$	$1/2$	$1/2$
$1/2$	0	0
0	1	0
0	$1/2$	0
0	0	0

α	β	$\alpha \vee \beta$
1	1	1
1	$1/2$	1
1	0	1
$1/2$	1	1
$1/2$	$1/2$	$1/2$
$1/2$	0	$1/2$
0	1	1
0	$1/2$	$1/2$
0	0	0

α	β	$\alpha \rightarrow \beta$
1	1	1
1	$1/2$	$1/2$
1	0	0
$1/2$	1	1
$1/2$	$1/2$	$1/2$
$1/2$	0	0
0	1	1
0	$1/2$	1
0	0	1

Figura 7 – Tabelas de verdade dos conectivos de LFI1.

Outros conectivos consagrados análogos aos clássicos podem ser facilmente definidos:

Definição 4.2.1 (Outros conectivos de LFI1)**.**

$$\sim\alpha := \neg\alpha \wedge \neg\bullet\alpha$$
$$\alpha \leftrightarrow \beta := (\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)$$
$$\circ\alpha := \neg\bullet\alpha$$

Agora é possível entender melhor a equivalência formal entre J3 e LFI1: basta observar que J3-fórmulas no formato $\nabla\alpha$ equivalem a LFI1-fórmulas do tipo $\alpha \vee \bullet\alpha$; analogamente, $\bullet\alpha \leftrightarrow \nabla\alpha \wedge \nabla\neg\alpha$ (CARNIELLI; MARCOS; DE AMO, 2004, p. 126). Por fim, adota-se a seguinte axiomática para LFI1:

Definição 4.2.2 (Axiomática de LFI1 (CARNIELLI; MARCOS; DE AMO, 2004, p. 131–132))**.**

i) $\alpha \rightarrow (\beta \rightarrow \alpha)$
ii) $(\alpha \rightarrow \beta) \rightarrow ((\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow (\alpha \rightarrow \gamma))$
iii) $\alpha \rightarrow (\beta \rightarrow (\alpha \wedge \beta))$
iv) $(\alpha \wedge \beta) \rightarrow \alpha$
v) $(\alpha \wedge \beta) \rightarrow \beta$
vi) $\alpha \rightarrow (\alpha \vee \beta)$
vii) $\beta \rightarrow (\alpha \vee \beta)$
viii) $(\alpha \rightarrow \gamma) \rightarrow ((\beta \rightarrow \gamma) \rightarrow ((\alpha \vee \beta) \rightarrow \gamma))$

[4] O conjunto minimal de conectivos adequados de LFI1 é $\{\neg, \bullet, \vee\}$, mas optou-se aqui por já partir de sua "assinatura completa" em parte por questões de espaço, em parte porque o enfoque desta dissertação não é, afinal, demonstrar resultados em LFI1 já conhecidos na literatura.

ix) $\alpha \vee \neg\alpha$

x) $\neg\neg\alpha \leftrightarrow \alpha$

xi) $\circ\alpha \rightarrow (\alpha \rightarrow (\neg\alpha \rightarrow \beta))$

xii) $\bullet\alpha \rightarrow (\alpha \wedge \neg\alpha)$

xiii) $\bullet(\alpha \wedge \beta) \leftrightarrow ((\bullet\alpha \wedge \beta) \vee (\bullet\beta \wedge \alpha))$

xiv) $\bullet(\alpha \vee \beta) \leftrightarrow ((\bullet\alpha \wedge \neg\beta) \vee (\bullet\beta \wedge \neg\alpha))$

xv) $\bullet(\alpha \rightarrow \beta) \leftrightarrow (\alpha \wedge \bullet\beta)$

Regra de inferência: $\{\alpha \rightarrow \beta,\ \alpha\} \models_{\text{LFI1}} \beta$ (*modus ponens*)

Teorema 4.2.1 (Completude de LFI1)**.** *Para toda teoria* $\Gamma \cup \{\varphi\}$ *de LFI1,* $\Gamma \vdash_{\text{LFI1}} \varphi \iff \Gamma \models_{\text{LFI1}} \varphi$.

Demonstração. Trata-se de uma prova extensa e complexa; os interessados em acompanhá-la são remetidos a Carnielli, Marcos e de Amo (2004, p. 134–135). ■

5 Probabilidade do condicional, probabilidade condicional, e paraconsistência

A demonstração da impossibilidade de uma tese é um dos gêneros de prova mais fortes nas ciências formais. De fato, comprovar que um resultado não pode ser obtido dentro de um sistema formal tem repercussões mais profundas, muitas vezes, que demonstrar um teorema desse sistema. Na Matemática, três problemas geométricos antigos provaram-se insolúveis no século XIX[1]: i) o problema da quadratura do círculo[2] (demonstrado insolúvel em 1882), ii) o problema de dobrar o volume de um cubo (demonstrado insolúvel em 1837), e iii) o problema de trissectar um ângulo (demonstrado insolúvel também em 1837). Na Lógica, os casos mais célebres desse gênero de metarresultado são, talvez, os teoremas da incompletude de Gödel, aludidos na seção 2.3. Nessa toada, dada uma versão de LFI1 equipada com uma semântica probabilística, seria desejável ou demonstrar resultados equivalentes, nesse sistema, às leis do excesso e aos teoremas de trivialização expostos no capítulo 3, ou demonstrar a impossibilidade de fazer isso. Embora nenhum desses resultados mais robustos tenha sido obtido, pode-se inferir, da observação cuidadosa das demonstrações do capítulo 3, que alguns de seus passos são interditados pelas propriedades formais de LFI1, de modo que uma demonstração de resultados análogos para esse sistema, caso exista, exigirá outras estratégias. Este capítulo pretende comentar a importância formal de cada uma das interdições mais relevantes.

5.1 Axiomas e definições

Os axiomas da Definição 3.0.1 adotados para as demonstrações das leis do excesso e dos resultados de trivialização são facilmente transponíveis para LFI1, porque não dependem de nenhuma propriedade intrinsecamente clássica. O axioma Cdmn apenas fixa o contradomínio das funções de probabilidade, que substituem, nessa semântica, as valorações tradicionais, como o intervalo real: $0 \leq p(\alpha) \leq 1$, para toda proposição α. Sendo uma regra meramente aritmética, ele pode ser empregado também em LFI1 *ipsis litteris*. O axioma Equi estipula que proposições logicamente equivalentes (isto é, tais que assumir uma como premissa basta para

---

[1] É importante notar que a formulação clássica desses problemas requer que eles sejam resolvidos com apenas régua e compasso idealizados, e em uma sequência finita de passos. Isto é, a régua deve ser concebida como um objeto com exatamente uma extremidade retilínea não graduada e que só existe enquanto em contato com a superfície em que as figuras são construídas. Em particular, esses objetos ideais não podem ser usados para medir e "transferir" comprimentos. Alguns dos problemas são solúveis caso essa restrição seja desconsiderada (caso a régua seja graduada, por exemplo).

[2] O problema consiste em construir um quadrado com a mesma área de um dado círculo.

inferir a outra) têm valores de probabilidade iguais: se $\alpha \dashv\models \beta$, então $p(\alpha) = p(\beta)$. A única adaptação necessária para a utilização desse axioma em LFI1 é que a condição "se $\alpha \dashv\models \beta$" deve ser entendida como abreviação de "se $\alpha \models_{\text{LFI1}} \beta$ e $\beta \models_{\text{LFI1}} \alpha$". O axioma Disj determina que a probabilidade da disjunção de duas proposições mutuamente excludentes é a soma das probabilidades individuais: se $\varnothing \models {\sim}(\alpha \wedge \beta)$, então $p(\alpha \vee \beta) = p(\alpha) + p(\beta)$. As contradições de LFI1 diferem das clássicas; o exemplo mais óbvio são as conjunções de uma proposição com sua negação, talvez as contradições clássicas mais intuitivas, e que são proposições notoriamente contingentes nas lógicas paraconsistentes. Ainda assim, como esse axioma não requer nenhuma propriedade específica das proposições cuja conjunção deve ser uma partícula *bottom*, a única adaptação necessária para seu uso em LFI1 é convencionar que $\dashv\models$ diz respeito a essa lógica. Por fim, o axioma Need estabelece que partículas *top* têm probabilidade máxima: se $\varnothing \models \alpha$, então $p(\alpha) = 1$. Novamente, as partículas *top* de LFI1 não coincidem totalmente com as de LPC, mas nada é dito sobre o conteúdo ou a forma de α exceto que ela deve ser tautológica.

Quanto às definições, as três necessárias para os resultados de excesso e trivialização analisados são a definição de probabilidade condicional (Definição 3.0.2), a definição de condicionalização (Definição 3.0.3), e a definição de independência (Definição 3.0.4). A primeira convenciona que $p(\alpha \mid \beta) := \frac{p(\alpha \wedge \beta)}{p(\beta)}$, se $p(\beta) > 0$. A segunda, que $p_{\beta}(\alpha) := p(\alpha \mid \beta)$, se $p(\beta) > 0$. E a terceira, que duas proposições α e β são independentes se, e somente se, $p(\alpha|\beta) = p(\alpha)$. Além da divisão (a operação aritmética), essas definições dependem apenas de propriedades da conjunção clássica preservadas por LFI1, como a comutatividade e a associatividade (Teoremas 2.1.1 e 2.1.2). A expectativa de que a conjunção dessa lógica tenha essas duas propriedades confirma-se, como atesta a Tabela 8.

α	β	γ	$\alpha \wedge \beta$	$\beta \wedge \alpha$	$\alpha \wedge (\beta \wedge \gamma)$	$(\alpha \wedge \beta) \wedge \gamma$
1	1	1	1	1	1	1
1	1	½	½	½	½	½
1	1	0	0	0	0	0
1	½	1	½	½	½	½
1	½	½	½	½	½	½
1	½	0	0	0	0	0
1	0	1	0	0	0	0
1	0	½	0	0	0	0
1	0	0	0	0	0	0
½	1	1	1	1	½	½
½	1	½	½	½	½	½
½	1	0	0	0	0	0
½	½	1	½	½	½	½
½	½	½	½	½	½	½
½	½	0	0	0	0	0

½	0	1	0	0	0	0
½	0	½	0	0	0	0
½	0	0	0	0	0	0
0	1	1	1	1	0	0
0	1	½	½	½	0	0
0	1	0	0	0	0	0
0	½	1	½	½	0	0
0	½	½	½	½	0	0
0	½	0	0	0	0	0
0	0	1	0	0	0	0
0	0	½	0	0	0	0
0	0	0	0	0	0	0

Tabela 8 – A tabela de verdade de LFI1 para os Teoremas 2.1.1 e 2.1.2.

A interpretação da conjunção de LFI1 é uma função $\wedge_{\text{LFI1}}: \{0, 1/2, 1\}^2 \mapsto \{0, 1/2, 1\}$ que associa cada par de valores de verdade ao menor dos dois. Isto é, poderia ser descrita como

$$\wedge_{\text{LFI1}}(i, j) := \begin{cases} i & \text{se } i \leq j \\ j & \text{do contrário} \end{cases}$$

A interpretação da disjunção de LFI1 é análoga: retorna o maior de seus dois argumentos[3]. É natural, portanto, que esses dois conectivos de LFI1 sejam comutativos e associativos, como os clássicos.

5.2 Lemas básicos

5.2.1 Lemas 3.1.1 e 3.1.2

O Lema 3.1.1 é o que estabelece que $p(\alpha \wedge \beta) = p(\alpha \mid \beta)p(\beta)$, se $p(\beta) > 0$, e o Lema 3.1.2 afirma que $p(\alpha \mid \beta) = p(\alpha)$ se, somente se, $p(\alpha \wedge \beta) = p(\alpha)p(\beta)$. Isto é, duas proposições são independentes (Definição 3.0.4) se, e somente se, a probabilidade absoluta de sua conjunção é igual ao produto das probabilidades individuais.

Esses dois lemas elementares são resultado direto da definição das funções de probabilidade condicional (Definição 3.0.2) e da definição de independência probabilística (Definição 3.0.4), e o único conectivo lógico que ocorre neles é a conjunção, que já se verificou ser "bem comportada" em LFI1 para os propósitos dos resultados do capítulo 3. Sendo assim,

3 Observe-se que essas duas funções são numéricas, portanto, diferentemente das funções de interpretação de LPC como esta foi axiomatizada e definida neste trabalho.

não surpreende que esses primeiros lemas possam ser facilmente transpostos para LFI1. Como seria de se esperar, os empecilhos a demonstrações em LFI1 análogas às de LPC começam a aparecer nos lemas em que a negação desempenha um papel fundamental.

5.2.2 Lema 3.1.3

O Lema 3.1.3 estabelece a relação aritmética entre $p(\alpha)$ e $p(\sim\alpha)$: $p(\alpha) = 1 - p(\sim\alpha)$. Sua demonstração emprega dois teoremas de LPC: $\alpha \vee \sim\alpha$ (Teorema 2.1.8), o chamado "princípio do terceiro excluído", e $\sim(\alpha \wedge \sim\alpha)$ (Teorema 2.1.10), o princípio da não-contradição (simples). A presença do número 1 torna conveniente recorrer ao axioma Need, e a subtração, sendo dual à adição, permite a aplicação do axioma Disj, por ser este o postulado em que ocorre essa operação aritmética. De fato, a demonstração começa (passo 3.8) invocando o princípio do terceiro excluído e o axioma Need, e prossegue (passo 3.9) lançando mão do princípio da não-contradição. O uso deste último princípio sugere que este lema não vale em LFI1, o que é intuitivamente esperado, já que a tolerância a contradições é justamente uma das principais características das lógicas paraconsistentes em geral. Com efeito, a última coluna da tabela de verdade abaixo (Tabela 9) confirma que a não-contradição não vale em LFI1: a valoração destacada em cinza atribui um valor distinguido a $\alpha \wedge \neg\alpha$, de modo que essa fbf não é uma partícula *bottom* de LFI1. Já a validade em LFI1 do princípio do terceiro excluído é menos óbvia à primeira vista, mas a penúltima coluna da mesma tabela evidencia sua validade também nessa lógica.

α	$\neg\alpha$	$\alpha \vee \neg\alpha$	$\alpha \wedge \neg\alpha$
1	0	1	0
½	½	½	½
0	1	1	0

Tabela 9 – A tabela de verdade dos princípios do terceiro excluído e da não-contradição (simples) em LFI1.

5.2.3 Lema 3.1.4 e Corolário 3.1.5

O Lema 3.1.4 enuncia que $p(\alpha) = p(\alpha \wedge \beta) + p(\alpha \wedge \sim\beta)$. Ele é particularmente útil porque relaciona a probabilidade absoluta de uma proposição à de outra. Trata-se de uma adaptação simples da equivalência clássica $\alpha \leftrightarrow (\alpha \wedge \beta) \vee (\alpha \wedge \sim\beta)$ (Teorema 2.1.9) à semântica probabilística. O primeiro passo da demonstração resulta justamente da aplicação do axioma Equi a essa equivalência. Uma vez que essa equivalência está diretamente ligada ao princípio do terceiro excluído (Teorema 2.1.8), e este vigora também em LFI1, como já comentado, seria de se esperar que ela também vigorasse. Essa expectativa confirma-se, como atesta a tabela de verdade abaixo.

α	β	$\alpha \wedge \beta$	$\alpha \wedge \neg\beta$	$\begin{bmatrix}(\alpha \wedge \beta) \vee \\ (\alpha \wedge \neg\beta)\end{bmatrix}$	$\alpha \rightarrow \begin{bmatrix}(\alpha \wedge \beta) \vee \\ (\alpha \wedge \neg\beta)\end{bmatrix}$	$\begin{bmatrix}(\alpha \wedge \beta) \vee \\ (\alpha \wedge \neg\beta)\end{bmatrix} \rightarrow \alpha$	$\alpha \leftrightarrow \begin{bmatrix}(\alpha \wedge \beta) \vee \\ (\alpha \wedge \neg\beta)\end{bmatrix}$
1	1	1	0	1	1	1	1
1	½	½	½	½	½	1	½
1	0	0	1	1	1	1	1
½	1	½	0	½	½	½	½
½	½	½	½	½	½	½	½
½	0	0	½	½	½	½	½
0	1	0	0	0	1	1	1
0	½	0	0	0	1	1	1
0	0	0	0	0	1	1	1

Tabela 10 – A tabela de verdade de LFI1 para o Lema 3.1.4.

O segundo passo da demonstração do Lema 3.1.4 baseia-se na contraditoriedade clássica de expressões da forma $(\alpha \wedge \beta) \wedge (\alpha \wedge \sim\beta)$ (Teorema 2.1.11). A expectativa é de que essas fbfs não sejam contradições de LFI1, uma vez que são “versões complexas” do princípio de não-contradição (Teorema 2.1.10), que não é teorema de LFI1. Essa expectativa confirma-se, como atesta a tabela de verdade abaixo. As valorações destacadas em cinza são as responsáveis por $(\alpha \wedge \beta) \wedge (\alpha \wedge \neg\beta)$ não ser uma contradição de LFI1.

α	β	$\neg\beta$	$\alpha \wedge \beta$	$\alpha \wedge \neg\beta$	$(\alpha\wedge\beta)\wedge(\alpha\wedge\neg\beta)$
1	1	0	1	0	0
1	½	½	½	½	½
1	0	1	0	1	0
½	1	0	½	0	0
½	½	½	½	½	½
½	0	1	0	½	0
0	1	0	0	0	0
0	½	½	0	0	0
0	0	1	0	0	0

Tabela 11 – A tabela de verdade de LFI1 para o Teorema 2.1.11.

O Corolário 3.1.5 é a contraparte do axioma Need, e estipula que partículas *bottom* tem probabilidade nula: $p(\bot) = 0$. Como naquele contexto, o conjunto das partículas *bottom* de LFI1 tem uma intersecção não-vazia com o respectivo conjunto de LPC, mas não coincide com este último; note-se que nada é dito sobre o conteúdo ou a forma de α exceto que ela deve ser receber um valor de verdade não-distinguido de toda valoração.

5.2.4 Lemas 3.1.6 e 3.1.7

O Lema 3.1.6 enuncia que, se $\beta \dashv\vDash \gamma$, então $p(\alpha \mid \beta) = p(\alpha \mid \gamma)$. Não é um resultado particularmente importante, e foi demonstrado apenas porque ocorre em várias outras provas, de modo que é conveniente dedicar-lhe um lema. Com efeito, ele baseia-se apenas na definição de probabilidade condicional (Definição 3.0.2) e no axioma Equi, que já foram constatados transponíveis para LFI1. Assim, pode ser empregado também nessa lógica.

O Lema 3.1.7 mostra que a probabilidade de uma proposição α relativa a uma tautologia é igual à probabilidade absoluta de α: $p(\alpha \mid \top) = p(\alpha)$. Além dos axiomas Equi e Need e da Definição 3.0.2, a demonstração desse Lema baseia-se no Teorema 2.1.4, segundo o qual uma conjunção em que uma das proposições é tautológica equivale, em LPC, à outra proposição: $\alpha \wedge \top \leftrightarrow \alpha$. Poder-se-ia pensar que, ao ser colocada em conjunção com uma tautologia, uma proposição (possivelmente contingente) mantém seu valor de verdade. Como este resultado concerne apenas a conjunção, que sabidamente funciona de maneira muito semelhante tanto em LPC quanto em LFI1, é natural que vigore também nesta última lógica, como comprova a tabela abaixo.

α	$\top$	$\alpha \wedge \top$
1	1	1
½	1	½
0	1	0

Tabela 12 – A tabela de verdade de LFI1 para o Teorema 2.1.4.

5.3 Lemas intermediários

5.3.1 Lema 3.2.1

O Lema 3.2.1 revela que a probabilidade de uma proposição condicionada a uma conjunção em que ela mesma ocorre afirmada é sempre máxima: $p(\alpha \mid \alpha \wedge \beta) = p(\alpha \mid \beta \wedge \alpha) = 1$, se $p(\alpha \wedge \beta) > 0$. A demonstração recorre a alguns resultados já constatados válidos em LFI1: o axioma Equi, a definição de probabilidade condicional (Definição 3.0.2), e o Lema da probabilidade condicional equivalente (Lema 3.1.6).

5.3.2 Lema 3.2.2

O Lema 3.2.2 é muito útil, porque fornece o valor de probabilidade de qualquer disjunção: $p(\alpha \vee \beta) = p(\alpha) + p(\beta) - p(\alpha \wedge \beta)$, ao passo que o axioma Disj é mais restrito, pois aplica-se somente a disjunções de proposições mutuamente excludentes. A demonstração deste Lema baseia-se em alguns resultados já analisados: o Lema da expansão (Lema 3.1.4), a comutatividade da conjunção (Teorema 2.1.1), o axioma Equi, e o Lema do complemento (Lema 3.1.3).

A novidade na demonstração do Lema 3.2.2 é a ocorrência da lei de De Morgan da conjunção (Teorema 2.1.6), isto é, a equivalência $\alpha \wedge \beta \leftrightarrow {\sim}({\sim}\alpha \vee {\sim}\beta)$. Essa regra vale também em LFI1, como evidencia a Tabela 13; as colunas em cinza destacam os valores de verdade das duas proposições equivalentes.

α	β	$\neg\alpha$	$\neg\beta$	$\alpha \wedge \beta$	$\neg\alpha \vee \neg\beta$	$\neg(\neg\alpha \vee \neg\beta)$
1	1	0	0	1	0	1
1	½	0	½	½	½	½
1	0	0	1	0	1	0
½	1	½	0	½	½	½
½	½	½	½	½	½	½
½	0	½	1	0	1	0
0	1	1	0	0	1	0
0	½	1	½	0	1	0
0	0	1	1	0	1	0

Tabela 13 – A tabela de verdade de LFI1 para a lei de De Morgan da conjunção (Teorema 2.1.6).

Apesar disso, o Lema 3.2.2 não pode ser transposto diretamente para LFI1, porque sua demonstração recorre a dois resultados inválidos nessa lógica: o Lema do complemento (Lema 3.1.3) e o Lema da expansão (Lema 3.1.4).

5.3.3 Lema 3.2.3

O Lema 3.2.3, $p(\alpha \rightarrow \beta) = p(\sim\alpha \vee \beta)$, é consequência direta da equivalência clássica $(\alpha \rightarrow \beta) \leftrightarrow (\sim\alpha \vee \beta)$ (Teorema 2.1.5) e do lema da disjunção (Lema 3.2.2). Como os axiomas adotados (Definição 3.0.1) e a definição de probabilidade condicional (Definição 3.0.2) envolvem apenas a conjunção e a disjunção, é conveniente "converter" as implicações em outros conectivos. O Lema 3.2.2 já foi analisado, de modo que cabe investigar a validade do Teorema 2.1.5. A equivalência clássica não vale em LFI1 devido à valoração destacada em cinza na Tabela 14, mas a implicação pode ser expressa em termos da negação fraca, do operador de inconsistência e da disjunção: $(\alpha \rightarrow \beta) \leftrightarrow (\beta \vee \neg(\alpha \vee \bullet\alpha))$. Ou seja, o Lema 3.2.3 original não vigora em LFI1, mas pode ser adaptado a essa lógica.

α	β	$\neg\alpha \vee \beta$	$\alpha \rightarrow \beta$	$\beta \vee \neg(\alpha \vee \bullet\alpha)$
1	1	1	1	1
1	½	½	½	½
1	0	0	0	0
½	1	1	1	1
½	½	½	½	½
½	0	½	0	0
0	1	1	1	1
0	½	1	1	1
0	0	1	1	1

Tabela 14 – A implicação material expressa em termos de outros conectivos de LFI1.

5.3.4 Lema 3.2.4

O Lema 3.2.4 é a caracterização probabilística da equivalência material de LPC: se $p(\alpha \leftrightarrow \beta) = 1$, então $p(\alpha) = p(\beta)$. Na semântica clássica usual, se a equivalência entre duas proposições é tautológica, então ambas têm o mesmo valor de verdade em todas as valorações. Analogamente, nessa semântica probabilística, se a probabilidade de uma equivalência entre duas proposições é máxima, então ambas têm o mesmo valor de probabilidade para todas as funções de probabilidade.

A demonstração desse Lema recorre aos seguintes resultados: o axioma Equi, a expressabilidade $\leftrightarrow$ em termos da implicação material e da conjunção $((\alpha\leftrightarrow\beta)\leftrightarrow(\alpha\rightarrow\beta\wedge\beta\rightarrow\alpha))$, o Lema da implicação material (Lema 3.2.3), o Lema da expansão (Lema 3.1.4), a comutatividade da conjunção (Teorema 2.1.1), e o Lema do complemento (Lema 3.1.3).

A equivalência material reduz-se, também em LFI1, a uma conjunção de implicações materiais recíprocas, porque a implicação material pode ser expressa em termos da negação, do operador de inconsistência e da disjunção nessa lógica, como já comentado. Não obstante, a formulação original deste lema não vigora em LFI1, por razões semelhantes àquelas que interditam a validade do Lema da implicação (Lema 3.2.3): a dependência de sua demonstração aos Lemas 3.1.3 e 3.1.4.

5.3.5 Lema 3.2.5 e Corolários 3.2.6 e 3.2.7

O Lema 3.2.5 explicita que é nula a probabilidade de uma conjunção em que uma das proposições está negada, caso a afirmação desta preposição tenha probabilidade máxima. Isto é, se $p(\alpha) = 1$, então $p(\sim\alpha \wedge \beta) = p(\beta \wedge \sim\alpha) = 0$.

O Corolário 3.2.6 é bastante semelhante, e enuncia que, se uma proposição tem probabilidade nula, então sua conjunção com qualquer proposição também tem probabilidade nula: se $p(\alpha) = 0$, então $p(\alpha \wedge \beta) = p(\beta \wedge \alpha) = 0$.

O Corolário 3.2.7 afirma que, se uma proposição é contraditória, então sua conjunção com qualquer outra proposição tem probabilidade nula: se $\varnothing \models \sim\alpha$, então $p(\alpha\wedge\beta) = p(\beta\wedge\alpha) = 0$. Claramente, este resultado é apenas um caso particular do Corolário 3.2.6, uma vez que o conjunto de contradições está contido no de proposições de probabilidade nula[4]. Isto é, uma proposição contingente pode ter probabilidade nula ou não, mas toda contradição tem probabilidade nula (Corolário 3.1.5).

Esses três resultados são úteis porque, se uma das proposições de uma conjunção tem valor de probabilidade extremo, então pode-se determinar a probabilidade de sua conjunção com outra proposição, mesmo que nada se saiba sobre esta última. As demonstrações dos três recorrem apenas a resultados já estudados: os axiomas Cdmn e Equi, a comutatividade da conjunção (Teorema 2.1.1), e os lemas do complemento (Lema 3.1.3) e da expansão (Lema 3.1.4).

[4] Dados o axioma Need e o Lema 3.1.3.

Os empecilhos à validade desses três resultados em LFI1 são justamente esses últimos dois lemas, que, como visto, não vigoram nessa lógica (algo bastante natural, dada sua paraconsistência).

5.3.6 Lema 3.2.8

O Lema 3.2.8 explica que a probabilidade de uma conjunção em que uma das proposições tem probabilidade máxima é igual à probabilidade da outra proposição: se $p(\alpha) = 1$, então $p(\alpha \wedge \beta) = p(\beta) = p(\beta \wedge \alpha)$. Mais uma vez, a demonstração baseia-se em alguns dos resultados já analisados: o axioma Equi, a comutatividade da conjunção (Teorema 2.1.1), os lemas do complemento (Lema 3.1.3) e da expansão (Lema 3.1.4), o lema da disjunção (Lema 3.2.2), e a lei de De Morgan da disjunção (Teorema 2.1.6). Esta última aparece de forma ligeiramente diferente, no entanto. No passo (3.62), temos a igualdade $p(\alpha \wedge \beta) = p({\sim}[\alpha \wedge {\sim}\beta])$. No passo seguinte, lê-se que $p(\alpha \wedge \beta) = p({\sim}\alpha \vee \beta)$. Ou seja, essa passagem recorre, além da lei de De Morgan, à regra ${\sim}{\sim}\alpha \leftrightarrow \alpha$ (Teorema 2.1.12):

$$\begin{aligned} p(\alpha \wedge \beta) &= p({\sim}\alpha \vee {\sim}{\sim}\beta) \\ &= p({\sim}\alpha \vee \beta) \end{aligned}$$

Em LPC, negar uma proposição um número ímpar de vezes equivale a negá-la uma única vez, e negá-la um número par de vezes equivale a afirmá-la (não negá-la). A negação fraca ($\neg$) compartilha dessa propriedade, porque sua respectiva função de interpretação também é inversível. Isto é, se $\overline{\neg}\colon \{0, 1/2, 1\} \mapsto \{0, 1/2, 1\}$ denotar essa função, então $\overline{\neg}^{-1}(\overline{\neg}(i)) = i$; a Tabela 15 explicita esse fato.

α	$\neg\alpha$	$\neg\neg\alpha$
1	0	1
1/2	1/2	1/2
0	1	0

Tabela 15 – A tabela de verdade de LFI1 para o Teorema 2.1.12.

5.3.7 Lema 3.2.9

O Lema 3.2.9 enuncia que, se uma conjunção tem probabilidade máxima, então cada proposição que a compõe também a tem: se $p(\alpha \wedge \beta) = 1$, então $p(\alpha) = p(\beta) = 1$. Ele é particularmente útil porque determina a probabilidade de duas proposições a partir de uma única fbf, o que permite algumas deduções interessantes. A demonstração depende apenas de resultados já analisados: os axiomas Cdmn e Equi, a comutatividade da conjunção (Teorema 2.1.1), e o Lema da expansão (Lema 3.1.4). É exatamente a ocorrência deste último que inviabiliza essa demonstração desse resultado para LFI1.

5.3.8 Lema 3.2.10

O Lema 3.2.10 afirma que é nula a probabilidade de uma proposição α relativa a uma conjunção em que ela ocorre negada: $p(\alpha \mid {\sim}\alpha \wedge \beta) = p(\alpha \mid \beta \wedge {\sim}\alpha) = 0$. A demonstração emprega apenas resultados já analisados: o axioma Equi, a comutatividade e a associatividade da conjunção (Teoremas 2.1.1 e 2.1.2), o princípio de não-contradição (Teorema 2.1.10), a definição de probabilidade condicional (Definição 3.0.2), o Corolário 3.2.7, e o Lema de probabilidade condicional equivalente (Lema 3.1.6). Como esperado, a dependência ao princípio de não-contradição interdita essa demonstração no contexto de LFI1.

5.3.9 Teorema 3.2.11

O Teorema 3.2.11 expressa a probabilidade absoluta de uma proposição em termos de sua probabilidade relativa a outra proposição: $p(\alpha) = p(\alpha \mid \beta)p(\beta) + p(\alpha \mid {\sim}\beta)p({\sim}\beta)$, se $0 < p(\beta) < 1$. O próprio enunciado deste teorema sugere que ele é consequência direta do Lema da expansão (Lema 3.1.4). Os demais resultados dos quais esse resultado depende são os Lemas da conjunção (Lema 3.1.1) e do complemento (Lema 3.1.3). É exatamente o uso desses dois últimos resultados que impede a transposição direta dessa demonstração para LFI1.

5.4 Leis do excesso

5.4.1 Primeira lei do excesso (Teorema 3.3.1)

A primeira lei do excesso de Popper mostra que a probabilidade de uma implicação material pode ser maior que a respectiva probabilidade condicional: $p(\alpha \rightarrow \beta) \geq p(\beta \mid \alpha)$, se $p(\alpha) > 0$. Ou seja, é o primeiro abalo à forte intuição de que esses dois valores deveriam coincidir.

A demonstração da primeira lei recorre aos princípios do terceiro excluído (Teorema 2.1.8) e da não-contradição (Teorema 2.1.10), ao teorema da implicação material (Teorema 2.1.5), aos axiomas Equi e Disj, e aos Lemas da conjunção (Lema 3.1.1) e do complemento (Lema 3.1.3). Como atestam as ocorrências do princípio da não-contradição e do lema do complemento, essa demonstração baseia-se em propriedades da negação clássica que, como visto, não vigoram em LFI1, de modo que um resultado análogo nessa lógica, se existir, precisará ser demonstrado de outra maneira.

5.4.2 Segunda lei do excesso (Teorema 3.3.2)

A segunda lei do excesso especifica em quais circunstâncias a igualdade demonstrada falsa pela primeira lei de fato é o caso: $p(\alpha \rightarrow \beta) = p(\beta \mid \alpha)$ se, e somente se, $p({\sim}\alpha) = 0$ ou

$p(\beta \mid \alpha) = 1$. Observe-se que, se $p(\beta \mid \alpha) = 1$, então

$$
\begin{aligned}
& p(\beta \mid \alpha) = 1 && \text{hipótese} \\
& \frac{p(\beta \wedge \alpha)}{p(\alpha)} = 1 && \text{Definição 3.0.2} \\
& p(\beta \wedge \alpha) = p(\alpha) && \\
& p(\alpha \wedge \beta) = p(\alpha) && \text{Comutatividade da conjunção (Teorema 2.1.1)}
\end{aligned}
$$

Ou seja, a igualdade intuitiva entre a implicação material e a respectiva probabilidade condicional ocorre apenas quando ou $p(\sim\alpha) = 0$, ou as duas proposições são independentes (Definição 3.0.4). Tratam-se de "casos extremos", por assim dizer, uma situação muito distinta daquela imaginada inicialmente.

A demonstração da segunda lei baseia-se na comutatividade e na associatividade da conjunção (Teoremas 2.1.1 e 2.1.2), nos princípios do terceiro excluído (Teorema 2.1.8) e da não-contradição (Teorema 2.1.10), e no Teorema da implicação material (Teorema 2.1.5), na definição de probabilidade condicional (Definição 3.0.2), e nos lema do complemento (Lema 3.1.3) e da conjunção (Lema 3.1.1). Os obstáculos à transposição da segunda lei a LFI1 são os resultados que concernem mais diretamente a negação clássica: o princípio da não-contradição (Teorema 2.1.10) e o lema do complemento (Lema 3.1.3).

Uma tentativa de restaurar, de algum modo, a suposta igualdade entre a implicação material e a probabilidade condicional seria a de estipular um novo conectivo condicional, de funcionamento tão similar quanto possível ao do primeiro, precisamente para apreender logicamente essa noção de teoria da probabilidade (Definição 3.3.1). Sendo denotado aqui pelo símbolo $\rightsquigarrow$, ele é tal que:

$$p(\alpha \rightsquigarrow \beta) \coloneqq p(\beta \mid \alpha), \text{ se } p(\alpha) > 0$$

A natureza semântica não verofuncional desse conectivo faz com que sua interação com os demais conectivos seja um tanto quanto convoluta. É importante, portanto, demonstrar alguns lemas intermediários que descrevem melhor seu funcionamento e são necessários para a demonstração da terceira lei do excesso e dos resultados de trivialização posteriores.

5.4.3 Lemas adicionais para os demais resultados

O Lema da probabilidade condicional composta (Lema 3.3.3) enuncia que a probabilidade de uma implicação probabilística[5] condicionada a uma proposição γ é igual à probabilidade do consequente dessa implicação condicionado à conjunção das outras duas fbfs: se $p(\alpha \wedge \gamma) > 0$, então $p(\alpha \rightsquigarrow \beta \mid \gamma) = p(\beta \mid \alpha \wedge \gamma)$.

Esse lema é útil porque "dá mobilidade" ao conectivo $\rightsquigarrow$: permite que ele interaja com a função de probabilidade condicional $p(\cdot \mid \cdot)$—o que é bastante natural, dado que ele supostamente é apenas outra maneira de expressar essa função. A demonstração desse lema

[5] Isto é, de uma fbf cujo conectivo principal é $\rightsquigarrow$.

recorre aos seguintes resultados: a associatividade da conjunção (Teorema 2.1.2), o axioma Equi, a definição de condicionalização (Definição 3.0.3), a definição do conectivo $\rightsquigarrow$ (Definição 3.3.1), a definição de probabilidade condicional (Definição 3.0.2), e a contrapositiva do Corolário 3.2.6. Como já explicado, esse último resultado não é facilmente convertível para o sistema de LFI1, dificuldade que se propaga a esse lema.

Como visto, o Corolário 3.2.6 não vigora em LFI1, o que impossibilita essa demonstração do Lema 3.3.3 para essa lógica. Não se sabe se há outra demonstração possível. Se não houver, então eventuais demonstrações de resultados análogos aos analisados aqui para essa lógica, caso existam, podem ser consideravelmente complexas: sem o Lema 3.3.3, o comportamento do conectivo $\rightsquigarrow$ em LFI1 fica muito restrito, porque torna-se particularmente difícil, por exemplo, converter proposições em que ele ocorre em proposições em que não ocorre.

O Lema 3.3.4 pode ser concebido como uma adaptação do Lema 3.3.3 à operação de condicionalização. Ele enuncia que, se $p(\alpha \wedge \beta) > 0$, então $p_\beta(\alpha \rightsquigarrow \gamma) = p(\gamma \,|\, \alpha \wedge \beta) = p(\gamma \,|\, \beta \wedge \alpha)$. Novamente, sua relevância deve-se principalmente às conversões que permite entre valores de probabilidade de implicações probabilísticas obtidos por condicionalização (Definição 3.0.3), e os respectivos valores de probabilidade relativa.

Nessa demonstração ocorrem os seguintes resultados: a comutatividade e associatividade da conjunção (Teoremas 2.1.1 e 2.1.2), o axioma Equi, a definição do conectivo $\rightsquigarrow$ (Definição 3.3.1), as definições de probabilidade condicional (Definição 3.0.2) e de condicionalização (Definição 3.0.3), e o Lema de probabilidade condicional equivalente (Lema 3.1.6). Todas essas premissas do Lema 3.3.4 são válidas em LFI1. Isso é bastante inusitado, uma vez que o Lema 3.3.3 não parece demonstrável nessa lógica, e a operação de condicionalização (Definição 3.0.3) apenas define uma função de probabilidade unária a partir de uma binária. Por que o conectivo $\rightsquigarrow$ parece interagir bem com as funções unárias geradas por condicionalização, mas não com as respectivas funções binárias? Poder-se-ia hipotetizar que isso sugere que um resultado análogo ao Lema 3.3.4 pode ser demonstrado em LFI1, por meio de outra estratégia de prova. Mas nada pode ser afirmado sobre isso no momento.

O teorema de LPC conhecido como "regra da exportação/importação" estabelece uma relação entre a implicação material e a conjunção: $p(\alpha \rightarrow [\beta \rightarrow \gamma]) = p((\alpha \wedge \beta) \rightarrow \gamma)$. Como o nome sugere, o lema homônimo da semântica probabilística (Lema 3.3.5) é uma adaptação dessa regra clássica na qual o condicional probabilístico $\rightsquigarrow$ substitui a implicação material $\rightarrow$: $p(\alpha \rightsquigarrow [\beta \rightsquigarrow \gamma]) = p((\alpha \wedge \beta) \rightsquigarrow \gamma) = p((\beta \wedge \alpha) \rightsquigarrow \gamma)$.

Assim como os dois Lemas anteriores estipulavam condições nas quais o conectivo $\rightsquigarrow$ poderia ser convertido em expressões equivalentes isentas dele, esse lema confere-lhe mobilidade ao determinar em que circunstâncias ele pode ser substituído por uma conjunção (e vice-versa). Mas esse resultado é ainda mais importante que os primeiros, porque aqueles concerniam a relação do conectivo com funções de probabilidade, mas ele já é definido a partir delas. O lema da exportação/importação, por outro lado, conecta-o à conjunção, um conectivo "propriamente lógico", o que é especialmente importante para que o novo condicional interaja não só com funções de probabilidade, mas também com os operadores lógicos tradicionais. Com efeito, Fitelson (2015) demonstra que esse lema é indipensável para a demonstração dos resultados de trivialização mais fortes que podem existir, como será analisado adiante.

A demonstração dessa regra baseia-se na comutatividade da conjunção (Teoremas 2.1.1), nos lemas de probabilidade condicional equivalente (Lema 3.1.6) e de probabilidade condicional composta (Lema 3.3.3), na definição do conectivo $\rightsquigarrow$ (Definição 3.3.1), e na contrapositiva do Corolário 3.2.6. O balanço geral do lema de exportação/importação é similar ao do Lema 3.3.3: a dependência da demonstração do caso clássico ao Corolário 3.2.6, que não vale em LFI1, evidencia a necessidade de adaptá-la às particularidades dessa lógica. Se isso for impossível, ou seja, se não houver como demonstrar em LFI1 um resultado análogo ao lema de exportação/importação, então as demonstrações da terceira lei do excesso e dos resultados de trivialização provavelmente também o serão.

5.4.4 Terceira lei do excesso (Teorema 3.3.6) e Lema 3.3.7

A terceira lei do excesso (Teorema 3.3.6) consiste no fato de que o valor de probabilidade de uma implicação probabilística relativa à respectiva implicação material é 1: se $p(\alpha \land \beta) > 0$, então $p(\alpha \rightsquigarrow \beta \mid \alpha \to \beta) = 1$. Embora de significado talvez menos expressivo que as duas primeiras leis do excesso, ela é um resultado importante para a demonstração dos resultados de trivialização. Sua demonstração recorre ao axioma Equi, à definição do conectivo $\rightsquigarrow$ (Definição 3.3.1), ao lema de exportação/importação (Lema 3.3.5), ao teorema de *modus ponens* internalizado (Teorema 2.1.7), e ao lema da probabilidade condicional redundante (Lema 3.2.1). Mais uma vez, uma das premissas atua como obstáculo à transposição do resultado para LFI1; nesse caso, é o lema da exportação/importação (Lema 3.3.5). Resta apenas observar a ocorrência do teorema de *modus ponens* internalizado (Teorema 2.1.7): a Tabela 16 evidencia que ele pode ser diretamente transposto para LFI1.

α	β	$\alpha \land \beta$	$\alpha \to \beta$	$(\alpha \to \beta) \land \alpha$
1	1	1	1	1
1	$1/2$	$1/2$	$1/2$	$1/2$
1	0	0	0	0
$1/2$	1	$1/2$	1	$1/2$
$1/2$	$1/2$	$1/2$	$1/2$	$1/2$
$1/2$	0	0	0	0
0	1	0	1	0
0	$1/2$	0	1	0
0	0	0	1	0

Tabela 16 – A tabela de verdade do Teorema 2.1.7 em LFI1.

O Lema 3.3.7 enuncia que a probabilidade de uma implicação probabilística é sempre igual ou maior à da implicação material de mesmos antecedente e consequente: se $p(\alpha \wedge \beta) > 0$, $p(\alpha \rightsquigarrow \beta) \geq p(\alpha \to \beta)$.

Esse resultado é a recíproca da primeira lei do excesso (Teorema 3.3.1), com o porém de que esta última expressa $p(\alpha \rightsquigarrow \beta)$ como $p(\beta \mid \alpha)$. Como se pode imaginar, a combinação desses dois resultados acarreta na igualdade, para toda função de probabilidade, entre as proposições da forma $p(\alpha \rightsquigarrow \beta)$ e aquelas da forma $p(\alpha \to \beta)$, o que significa que a tentativa de expressar probabilidades condicionais como probabilidades de um conectivo lógico é mal sucedida, pelo menos no caso de LPC.

A demonstração desse Teorema depende do axioma Cdmn, do Lema da conjunção (Lema 3.1.1), do Lema da expansão (Lema 3.1.4), e da terceira lei do excesso (Teorema 3.3.6). Essas duas últimas premissas são as que exigem adaptações de uma demonstração desse resultado em LFI1, caso exista.

5.5 Teoremas de trivialização

5.5.1 Lewis (1976)

5.5.1.1 Primeiro resultado de trivialização de Lewis (Teorema 3.4.1) e Lema 3.4.2

O enunciado do primeiro resultado de trivialização de Lewis (1976) (Teorema 3.4.1) é que, se $p(\alpha \wedge \beta) > 0$ e $p(\alpha \wedge {\sim}\beta) > 0$, então $p(\alpha \rightsquigarrow \beta) = p(\beta)$. Pela definição do conectivo $\rightsquigarrow$ (Definição 3.3.1), isso significa que, se $p(\alpha \wedge \beta) > 0$ e $p(\alpha \wedge {\sim}\beta) > 0$, então $p(\beta \mid \alpha) = p(\beta)$, ou seja, β e α são proposições independentes (Definição 3.0.4). Como p pode ser qualquer função de probabilidade e β e α quaisquer proposições, esse resultado mostra que assumir a existência do conectivo $\rightsquigarrow$ permite deduzir a independência de um par arbitrário de proposições perante uma função de probabilidade também arbitrária. É precisamente esse o sentido da expressão "trivialização": trata-se de um resultado que torna inexpressiva a lógica em questão, que perde a capacidade de distinguir proposições quanto às suas probabilidades de serem verdadeiras. É um fenômeno análogo à "explosão" que ocorre quando uma contradição é admitida em uma lógica não paraconsistente.

A demonstração desse primeiro resultado de Lewis (1976) recorre à contrapositiva do Corolário 3.2.6, ao Lema da probabilidade condicional composta (Lema 3.3.3), ao Lema da probabilidade condicional redundante (Lema 3.2.1), ao Lema da probabilidade condicional contraditória (Lema 3.2.10), e ao Teorema da probabilidade total (Teorema 3.2.11). A avaliação dessa demonstração é a mais díspar entre os resultados analisados até aqui: apenas o Lema 3.2.1 pode ser empregado em LFI1 de maneira direta; todas as outras premissas ou exigiriam outra estratégia de demonstração, ou não podem ser demonstradas nessa lógica.

O Lema 3.4.2 enuncia que, se $p(\alpha \wedge \beta) > 0$, então $p(\alpha \rightsquigarrow \beta) = p(\alpha \to \beta)$. Trata-se de uma forma de explicitar o colapso do condicional probabilístico frente à implicação material clássica, de modo que ambos os conectivos estão sujeitos ao primeiro resultado de trivialização (e às leis do excesso).

Sua demonstração recorre ao axioma Cdmn, à definição do conectivo $\rightsquigarrow$ (Definição 3.3.1), aos lemas 3.1.1 e 3.3.7, e à primeira lei do excesso (Teorema 3.3.1). Com efeito, pode-se pensar que esse lema é simplesmente corolário do Lema 3.3.7 e da primeira lei do excesso (Teorema 3.3.1); como esses dois resultados não são diretamente transponíveis para LFI1, este lema também não o é.

5.5.1.2 Segundo resultado de trivialização de Lewis (Teorema 3.4.3) e Lema 3.4.5

O segundo resultado de trivialização de Lewis (1976) (Teorema 3.4.3) é o fato de que não podem haver, numa LPC equipada com a semântica probabilística e o conectivo $\rightsquigarrow$ (Definição 3.3.1) do capítulo 3, três proposições cada qual com probabilidade positiva, mas incompatíveis quando tomadas duas a duas. Ou seja, a linguagem do conectivo $\rightsquigarrow$ inevitavelmente "desperdiça" quase todo o ganho de expressividade supostamente obtido pela incorporação de Teoria da probabilidade à Lógica.

A demonstração desse resultado baseia-se nos axiomas Equi e Disj, nas Definições 3.0.2 e 3.3.1, e no primeiro resultado de Lewis (1976) (Teorema 3.4.1). Dado o que já foi exposto neste capítulo, esse segundo resultado vigorará em LFI1 se, e somente se, o primeiro resultado for demonstrável.

O Lema 3.4.5 consiste no fato de que, se $0 < p(\alpha) < 1$, então $p_\alpha(\beta) \in \{0, 1\}$. Ou seja, condicionalizar sobre uma proposição de valor de probabilidade intermediário resulta em uma função de probabilidade "degenerada", isto é, que reduz-se a uma valoração usual. Este lema é um passo importante na direção do último resultado de Lewis (1976).

Sua demonstração baseia-se nas definições de probabilidade condicional, de condicionalização, e do conectivo $\rightsquigarrow$ (Definições 3.0.2, 3.0.3 e 3.3.1), no lema do complemento (Lema 3.1.3), na segunda lei do excesso (Teorema 3.3.2), e no segundo resultado de trivialização de Lewis (1976) (Teorema 3.4.3). Esses três últimos resultados dificultariam uma eventual demonstração deste Lema em LFI1. Em particular, note-se que o segundo resultado de Lewis (1976) parece interdepender do primeiro, e este parece exigir muitas adaptações para ser provado em LFI1, de modo que o segundo resultado também deve requerer uma estratégia de demonstração muito diferente, se este for o caso.

5.5.1.3 Terceiro resultado de trivialização de Lewis (Teorema 3.4.6)

O terceiro resultado de trivialização de Lewis (1976) (Teorema 3.4.6) enuncia que, se $0 < p(\alpha) < 1$, então a função $p_\alpha(\beta)$ atribui no máximo quatro valores diferentes a seu argumento: $0, p({\sim}\alpha), p(\alpha)$, ou 1.

Essa demonstração depende das definições de condicionalização (Definição 3.0.3), do conectivo $\rightsquigarrow$ (Definição 3.3.1) e de probabilidade condicional (Definição 3.0.2), do lema do complemento (Lema 3.1.3), e do segundo resultado de trivialização de Lewis (1976) (Teorema 3.4.3). Todas as definições, por si mesmas, podem ser aplicadas a LFI1 sem adaptações. São o lema do complemento (Lema 3.1.3) e o segundo resultado de trivialização de Lewis (1976) (Teorema 3.4.3), como em outros resultados, os responsáveis pela necessidade de correções na demonstração desse terceiro resultado, se ele puder ser provado em LFI1.

5.5.2 Teoremas de trivialização de Fitelson (2015) (seção 3.4.2)

O método empregado por Fitelson (2015) é um tanto diferente dos de Popper e Lewis (1976). Ele foi descrito na seção 3.4.2, mas cabe recordar que baseia-se em enumerar, em uma tabela estocástica, todos os possíveis valores de probabilidade que as variáveis proposicionais relevantes para a proposição analisada recebem, e então associar cada função de probabilidade à respectiva cláusula, de tal modo que a disjunção dessas cláusulas seja a FND equivalente à proposição em questão. Fitelson (2015) prossegue, então, instanciando várias vezes o lema da probabilidade condicional composta (Lema 3.3.3), demonstrando que todas as linhas da tabela estocástica, exceto duas, têm probabilidade nula.

5.5.2.1 Teorema 3.4.7

O Teorema 3.4.7 enuncia que, uma vez assumida a definição do conectivo $\rightsquigarrow$ (Definição 3.3.1), o lema da exportação/importação (Lema 3.3.5) equivale ao lema da probabilidade condicional composta (Lema 3.3.3): se $p(\alpha \land \gamma) > 0$, é o caso que $p(\alpha \rightsquigarrow [\beta \rightsquigarrow \gamma]) = p((\alpha \land \beta) \rightsquigarrow \gamma) = p((\beta \land \alpha) \rightsquigarrow \gamma)$ se, e somente se, também é o caso que $p(\alpha \rightsquigarrow \beta \mid \gamma) = p(\beta \mid \alpha \land \gamma)$.

Essa equivalência é interessante por ser uma simplificação e generalização dos resultados de trivialização originais de Lewis (1976). Isto é, Fitelson (2015) demonstra que, para obter os resultados mais fortes dessa família, bastam algumas poucas assunções muito elementares, a saber: a comutatividade da conjunção (Teorema 2.1.1), a definição de $\rightsquigarrow$ (Definição 3.3.1), o lema de probabilidade condicional equivalente (Lema 3.1.6), e um dos seguintes lemas: o lema da exportação/importação (Lema 3.3.5), ou o lema da probabilidade condicional composta (Lema 3.3.3). O outro lema segue-se apenas dessas premissas.

5.5.2.2 Lemas 3.4.8–3.4.10

As várias instâncias do lema da probabilidade condicional composta (Lema 3.3.3) usam as seguintes premissas oriundas ou de LPC, ou das definições mais básicas da semântica probabilística: o teorema do *modus ponens* internalizado (Teorema 2.1.7), o axioma Cdmn, e a definição de probabilidade condicional (Definição 3.0.2). Além disso, são empregados os seguintes lemas:

Lema 3.3.3 probabilidade condicional composta,

Lema 3.2.10 probabilidade condicional contraditória,

Lema 3.1.6 probabilidade condicional equivalente,

Lema 3.2.1 probabilidade condicional redundante, e

Lema 3.1.7 conjunção tautológica.

Os empecilhos à conversão da demonstração original ao sistema de LFI1 são os lemas da probabilidade condicional composta (Lema 3.3.3), da probabilidade condicional contraditória (Lema 3.2.10), e da conjunção tautológica (Lema 3.2.10).

6 Considerações finais

Now is better than never.
Although never is often better than *right* now.

Tim Peters, *The Zen of Python*.

Esta dissertação procurou investigar a viabilidade de demonstrarem-se resultados análogos às chamadas leis do excesso de Popper e aos resultados de trivialização de Lewis (1976) na lógica paraconsistente LFI1. Os resultados originais concernem a suposta igualdade entre probabilidades condicionais, da forma $p(\beta \mid \alpha)$, e probabilidades de proposições clássicas condicionais, da forma $p(\alpha \rightsquigarrow \beta)$, o símbolo $\rightsquigarrow$ denotando um conectivo de funcionamento lógico razoavelmente semelhante ao da implicação material. Posto de outro modo, esses resultados são demonstrações da impossibilidade da conexão aparentemente muito natural entre um conceito da Teoria da probabilidade e um conceito da Lógica (clássica). Nesse caso clássico, essa impossibilidade consiste no fato de que supor a igualdade $p(\alpha \rightsquigarrow \beta) = p(\beta \mid \alpha)$ implica na trivialização da função de probabilidade p. As demonstrações analisadas nesta dissertação foram as leis de excesso de Popper como expostas em Dorn (1992–93), e os resultados de trivialização de Lewis (1976) (à luz de Milne (2003)) e de Fitelson (2015).

Dado esse quadro para o ânbito clássico, as perguntas que motivaram esta pesquisa foram: quais as consequências, para esses resultados, de substituir LPC por uma lógica paraconsistente, em particular, pelo sistema LFI1? Teoremas análogos podem ser obtidos nessa lógica? Se sim, em que medida e por quais razões as estratégias de demonstração divergiriam daquelas empregadas por Popper, Lewis (1976) e Fitelson (2015)? Se não, seria possível obter um resultado negativo, isto é, uma prova de que resultados como esses *não* podem ser teoremas de LFI1?

Foi possível constatar que, de modo geral, a negação clássica desempenha papel fundamental nos resultados estudados, motivo pelo qual alguns lemas importantes não puderam ser reproduzidos em LFI1. Pode-se especular que a negação mais fraca dessa lógica interdita resultados como esses. Mas é possível que tais teoremas apenas exijam demonstrações diferentes, possivelmente mais complexas ou convolutas que as clássicas. Ou seja, o fato de que os resultados estudados não foram obtidos aqui para LFI1 não é um resultado de impossibilidade; representa apenas um primeiro passo na linha de investigação seguida aqui. Seria certamente desejável ou demonstrar resultados análogos em LFI1, ou obter uma prova da impossibilidade de fazê-lo, mas esses são desenvolvimentos para pesquisas futuras. Uma possível tática para atingir esse objetivo é o uso de raciocínios análogos aos utilizados na obtenção de paradoxos em lógica positiva.

O tema abordado nesta dissertação propiciou também uma oportunidade de discutir certos problemas filosóficos inerentes às interpretações da teoria de probabilidade, algo relevante na medida em que novas teorias de probabilidade, baseadas em lógicas não clássicas, constituem uma frente de pesquisa bastante promissora. O debate acerca da interpretação pode ser recolocado nesse novo contexto no sentido de que raciocínios probabilísticos admitem informações contraditórias, por exemplo. A relação entre lógica paraconsistente e teoria da probabilidade é bastante recente—de certo modo, assim como o desenvolvimento da própria Teoria da probabilidade, como comentado no capítulo 1. Ou seja, há muito a fazer nesse campo de pesquisa, inclusive quanto às suas possíveis aplicações práticas, assunto que fugiu ao escopo desta dissertação; restringimo-nos, aqui, a certos aspectos teóricos e conceituais nesse âmbito.

Referências

BUENO-SOLER, J. **Semântica algébrica de traduções possíveis**. 167 p. Dissertação (Mestrado) — Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas, Campinas, 2004. Disponível em: <http://www.repositorio.unicamp.br/handle/REPOSIP/279780>. Acesso em: 12 dez. 2015.

BUENO-SOLER, J.; CARNIELLI, W. A. May be and may be not: paraconsistent probabilities from the LFI viewpoint. **CLE e-prints**, v. 15, n. 2, 2015. Disponível em: <https://www.cle.unicamp.br/eprints/index.php/CLE_e-Prints/article/view/985>. Acesso em: 20 ago. 2018.

____. Paraconsistent probabilities: consistency, contradictions and Bayes' theorem. **Entropy**, v. 18, n. 9, 2016. Disponível em: <http://www.mdpi.com/1099-4300/18/9/325>. Acesso em: 21 jan 2017.

CARNIELLI, W. A.; CONIGLIO, M. E. **Paraconsistent logic**: Consistency, contradiction and negation. Springer International Publishing, 2016. v. 40. (Logic, Epistemology, and the Unity of Science, v. 40). ISBN 978-3-319-33205-5. Disponível em: <https://www.springer.com/br/book/9783319332031>. Acesso em: 8 nov. 2018.

CARNIELLI, W. A.; MARCOS, J.; DE AMO, S. Formal inconsistency and evolutionary databases. **Logic and logical Philosophy**, v. 8, n. 8, p. 115–152, 2004. ISSN 2300-9802. Disponível em: <https://apcz.umk.pl/czasopisma/index.php/LLP/article/view/LLP.2000.008>. Acesso em: 28 out. 2019.

CLIMENHAGA, N. The concept of probability is not as simple as you think. **Aeon**, 2019. Disponível em: <https://aeon.co/ideas/the-concept-of-probability-is-not-as-simple-as-you-think>. Acesso em: 6 maio 2019.

DE AMO, S.; CARNIELLI, W. A.; MARCOS, J. A logical framework for integrating inconsistent information in multiple databases. In: EITER, T.; SCHEWE, K.-D. (Ed.). **Foundations of Information and Knowledge Systems**. Berlim, Heidelberg: Springer Berlin Heidelberg, 2002. p. 67–84. ISBN 978-3-540-45758-9. Disponível em: <https://link.springer.com/chapter/10.1007%2F3-540-45758-5_5>. Acesso em: 29 out. 2019.

DEVLIN, K. **The unfinished game**: Pascal, Fermat, and the seventeenth-century letter that made the world modern. Nova York: Basic Books, 2008.

DORN, G. J. W. Popper's laws of the excess of the probability of the conditional over the conditional probability. **Conceptus: Zeitschrift für Philosophie**, v. 26, n. 67, p. 3–61, 1992–93. Disponível em: <https://philpapers.org/rec/DORPLO>. Acesso em: 18 jul. 2018.

FITELSON, B. A decision procedure for probability calculus with applications. **The Review of Symbolic Logic**, Cambridge University Press, v. 1, n. 1, p. 111–125, 2008. Disponível em: <https://doi.org/10.1017/S1755020308080039>. Acesso em: 14 jan. 2019.

____. The strongest possible Lewisian triviality result. **Thought: a journal of Philosophy**, v. 4, n. 2, p. 69–74, 2015. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/tht3.159>. Acesso em: 26 maio 2018.

GALAVOTTI, M. C. The interpretation of probability: still an open issue? **Philosophies**, v. 2, n. 3, 2017. ISSN 2409-9287. Disponível em: <http://www.mdpi.com/2409-9287/2/3/20>. Acesso em: 7 maio 2019.

GILLIES, D. **Philosophical theories of probability**. Londres, Nova York: Routledge, 2000.

____. Varieties of propensity. **The British Journal for the Philosophy of Science**, v. 51, n. 4, p. 807–835, 12 2000. ISSN 0007-0882. Disponível em: <https://doi.org/10.1093/bjps/51.4.807>. Acesso em: 7 out. 2019.

GOMES, E. L. **Sobre a história da paraconsistência e a obra de da Costa: a instauração da lógica paraconsistente**. 666 p. Tese (Doutorado) — Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, 2013. Disponível em: <http://repositorio.unicamp.br/jspui/handle/REPOSIP/281158>. Acesso em: 24 nov. 2018.

GOOD, I. J. Kinds of probability. **Science**, American Association for the Advancement of Science, v. 129, n. 3347, p. 443–447, 1959. ISSN 0036-8075. Disponível em: <http://science.sciencemag.org/content/129/3347/443>. Acesso em: 13 nov. 2018.

HAACK, S. **Philosophy of logics**. Cambridge: Cambridge University Press, 1978.

HACKING, I. **The emergence of probability**: A philosophical study of early ideas about probability induction and statistical inference. 2. ed. Nova York: Cambridge University Press, 2006.

HÁJEK, A. Interpretations of probability. In: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Winter 2012. Metaphysics Research Lab, Stanford University, 2012. Disponível em: <https://plato.stanford.edu/archives/win2012/entries/probability-interpret/>. Acesso em: 11 maio 2019.

KOLMOGOROV, A. N. **Foundations of the theory of probability**. Tradução de Nathan Morrison. Nova York: Chelsea Publishing Company, 1950. Disponível em: <https://archive.org/details/foundationsofthe00kolm>. Acesso em: 24 abr. 2018.

LAPLACE, P.-S. **A philosophical essay on probabilities**. Tradução de F. W. Truscott e F. L. Emory. Nova York: Cosimo, 2007.

LEWIS, D. Probabilities of conditionals and conditional probabilities. **The Philosophical Review**, v. 85, n. 3, p. 297–315, jul. 1976. Disponível em: <http://www.jstor.org/stable/2184045>. Acesso em: 20 abr. 2018.

MARCOS, J. **Logics of formal inconsistency**. 378 p. Tese (Doutorado) — Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas; Instituto Superior Técnico, Universidade Técnica de Lisboa, Campinas, 2005. Disponível em: <https://www.math.tecnico.ulisboa.pt/~jmarcos/Thesis/>. Acesso em: 24 nov. 2018.

MARCOS, J. (Wittgenstein & paraconsistência). **Principia: an international journal of epistemology**, v. 14, n. 1, p. 135–73, 2010. ISSN 1808-1711. Disponível em: <https://periodicos.ufsc.br/index.php/principia/article/view/1808-1711.2010v14n1p135>. Acesso em: 24 nov. 2018.

MENDELSON, E. **Introduction to mathematical Logic**. 6. ed. Boca Raton: Chapman and Hall/CRC Press, 2015. (Textbooks in Mathematics).

MILNE, P. The simplest Lewis-style triviality proof yet? **Analysis**, v. 63, n. 4, p. 300–303, oct 2003. Disponível em: <http://www.jstor.org/stable/3329270>. Acesso em: 24 abr. 2018.

POPPER, K. R. The propensity interpretation of the calculus of probability and of the quantum theory. In: KÖRNER, S.; PRYCE, M. H. L. (Ed.). **Observation and interpretation in the philosophy of Physics**: with special reference to Quantum Mechanics. Nova York: Dover, 1957, (Colston papers, v. 9).

____. The propensity interpretation of probability. **The British Journal for the Philosophy of Science**, X, n. 37, p. 25–42, 05 1959. ISSN 0007-0882. Disponível em: <https://doi.org/10.1093/bjps/X.37.25>. Acesso em: 1 jun. 2019.

PRETO, S. M. da S. **Lógica, probabilidade, consequência**. 109 p. Dissertação (Mestrado) — Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas, Campinas, 2015. Disponível em: <http://repositorio.unicamp.br/jspui/handle/REPOSIP/279772>. Acesso em: 19 ago. 2018.

ROEPER, P.; LEBLANC, H. **Probability theory and probability logic**. 1. ed. Toronto: University of Toronto Press, 1999. (Toronto Studies in Philosophy).

VINEBERG, S. Paradoxes of probability. In: BANDYOPADHYAY, P. S.; FORSTER, M. R. (Ed.). **Philosophy of Statistics**. Amsterdã: North-Holland, 2011, (Handbook of the Philosophy of Science, v. 7). p. 713–736. Disponível em: <http://www.sciencedirect.com/science/article/pii/B9780444518620500228>. Acesso em: 9 abr. 2019.